NAVAL POSTGRADUATE SCHOOL
MONTEREY, CALIFORNIA

THESIS

FITTING LANCHESTER AND OTHER EQUATIONS TO THE BATTLE OF KURSK DATA

By

Turker Turkes

March 2000

Thesis Advisor:	Thomas W. Lucas
Second Reader:	Ron Fricker

Approved for public release; distribution is unlimited.

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE March 2000	3. REPORT TYPE AND DATES COVERED Master's Thesis

4. TITLE AND SUBTITLE
Fitting Lanchester and other equations to the Battle of Kursk Data

5. FUNDING NUMBERS

6. AUTHOR(S)
Turkes, Turker

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Postgraduate School Monterey, CA 93943-5000

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
The views expressed in this thesis are those of the authors and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(maximum 200 words)*

This thesis extends previous research on validating Lanchester's equations with real data. The quality of the available historical data for validation of attrition models is poor. Most accessible battle data contain only starting sizes and casualties, sometimes only for one side. A detailed database of the Battle of Kursk of World War II, the largest tank battle in history, has recently been developed. The data were collected from military archives in Germany and Russia by the Dupuy Institute (TDI) and were reformatted into a computerized data base, designated as the Kursk Data Base (KDB), and recently made available and documented in the KOSAVE (Kursk Operation Simulation and Validation Exercise of the US Army) study. The data are two-sided, time phased (daily), and highly detailed. They cover 15 days of the campaign. This thesis examines how the various derivatives of Lanchester's equations fit the newly compiled database on the Battle of Kursk. In addition, other functional forms are fit. These results are contrasted with earlier studies on the Ardennes campaign. It turns out that a wide variety of models fit the data about as well. Unfortunately, none of the basic Lanchester models fit the data, bringing into question their use in combat modeling.

14. SUBJECT TERMS
Combat Modeling, Lanchester Equations, Battle of Kursk

15. NUMBER OF PAGES

15. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18 298-102

Approved for public release; distribution is unlimited

FITTING LANCHESTER AND OTHER EQUATIONS TO THE BATTLE OF KURSK DATA.

Turker Turkes
First Lieutenant, Turkish Army
B.S., Turkish Army Academy, 1993

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
March 2000

Author: Turker Turkes

Approved by: Thomas W.Lucas, Thesis Advisor

Ron Fricker, Second Reader

for Matthew G. B— CDR, USN

Richard E. Rosenthal, Chairman,
Department of Operations Research

ABSTRACT

This thesis extends previous research on validating Lanchester's equations with real data. The quality of the available historical data for validation of attrition models is poor. Most accessible battle data contain only starting sizes and casualties, sometimes only for one side. A detailed database of the Battle of Kursk of World War II, the largest tank battle in history, has recently been developed. The data were collected from military archives in Germany and Russia by the Dupuy Institute (TDI) and were reformatted into a computerized data base, designated as the Kursk Data Base (KDB), and recently made available and documented in the KOSAVE (Kursk Operation Simulation and Validation Exercise of the US Army) study. The data are two-sided, time phased (daily), and highly detailed. They cover 15 days of the campaign. This thesis examines how the various derivatives of Lanchester's equations fit the newly compiled database on the Battle of Kursk. In addition, other functional forms are fit. These results are contrasted with earlier studies on the Ardennes campaign. It turns out that a wide variety of models fit the data about as well. Unfortunately, none of the basic Lanchester models fit the data, bringing into question their use in combat modeling.

TABLE OF CONTENTS

EXECUTIVE SUMMARY

War is a conflict between nations or states carried on by force of considerable duration and magnitude, by land, sea, or air for obtaining and establishing the superiority and dominion of one over the other for some cause. Throughout history, war has been a topic of analysis for scientists and researchers, especially following World War II.

Soviets argue that Osipov [Ref.2] was the first to study and discover the equations most often used when modeling attrition in combat. The equations are widely known as "Lanchester's equations." Regardless of claims of prior or parallel discovery, Lanchester's equations for attrition provided the origin for modeling attrition in the United States and around the world. Today, with the advent of computers, Lanchester-based models of warfare are widely used in the decisionmaking process for research, development, acquisition of weapons systems, force mix decisions, and for aiding in the development of operational plans.

The basic generalized Lanchester Equations are of the form [Ref.6]:

$$\dot{B}(t) = aR(t)^{p} B(t)^{q} \tag{1}$$

$$\dot{R}(t) = bB(t)^{p} R(t)^{q} \tag{2}$$

where $B(t)$ and $R(t)$ are the strengths of blue and red forces at time t, $\dot{B}(t)$ and $\dot{R}(t)$ are the rates at which blue and red force levels are changing at time t, a and b are attrition parameters, p is the exponent parameter of the attacking force, and q is the exponent parameter of the defending force.

Two versions of the Lanchester equations are of particular interest. When $p = q = 1$, force ratios remain equal if $aR(0) = bB(0)$, and hence this condition is called, *Lanchester's linear law*. The interpretation of Lanchester's linear law is that a battle

governed by this model is characterized as a collection of small engagements, and was proposed by Lanchester as a model for ancient warfare. The equation is also considered a good model for area fire weapons, such as artillery. Lanchester contrasted the Linear Law with the condition $p = 1$, $q = 0$, which is called *Lanchester's square law*, where the force ratios remain equal when $aR(0)^2 = bB(0)^2$. He theorized that the square law applies to modern warfare, in which both sides are able to aim their fire. His model suggests that in modern warfare combatants should concentrate their forces. A third version with $p = 0$, $q = 1$ is called *Lanchester's logarithmic law.*

Past empirical validation studies of Lanchester Equations include the work of Bracken on the Ardennes campaign of World War II, Fricker, also on the Ardennes campaign, Clemens on the Battle of Kursk of World War II, and Hartley and Helmbold on the Inchon-Seoul campaign of the Korean War. These works are among the few quantitative studies that use daily force size data for real battles.

Bracken formulated four different models for the Ardennes campaign, which are variations of the basic Lanchester equations, and estimated their parameters for the first ten days of the of the Ardennes campaign (December 15, 1944 through January 16, 1945). He concluded that: (1) the Lanchester linear model best fits the Ardennes campaign data, (2) when combat forces are considered, allied individual effectiveness is greater than German individual effectiveness, (3) when total forces are considered, individual effectiveness is the same for both sides, and (4) there is an attacker advantage throughout the campaign.

Fricker's paper revisited Bracken's modeling of the Ardennes campaign, using linear regression to fit the total body of data from the entire campaign, including air sortie

data. Fricker concludes by saying that one side's losses are more a function of his own forces than a function of the opponent's forces, like the logarithmic law, and gives the Gulf War as an example to support this theory.

Clemens' analysis examined the validity of the Lanchester Models as they are applied to modern warfare using data from the Battle of Kursk. His analysis is an extension of Bracken's and Fricker's analyses of the Ardennes Campaign. He concludes that the Lanchester logarithmic model in both scalar and matrix form fits better than the Lanchester linear and square models.

Hartley and Helmbold tested Lanchester's square law using the data from the Inchon-Seoul campaign. They conclude that: (1) the data do not fit a constant coefficient Lanchester square law, (2) the data better fit a set of three separate battles (one distinct battle every six or seven days), (3) Lanchester's square law is not a proven attrition algorithm for warfare (but neither can it be completely discounted), and (4) more real combat data are needed to validate any proposed attrition law.

This thesis takes a closer look at Lanchester's equations using recently available data on the battle of Kursk. In July 1943, the Battle of Kursk, the largest tank battle in history, took place around the city of Kursk, Russia, and ended in the defeat of the Germans. A detailed database of this battle was recently developed. The data were collected from military archives in Germany and Russia by the Dupuy Institiute (TDI), and are reformatted into a computerized data base, designated as the Kursk Data Base (KDB). KDB is recently documented in the KOSAVE (Kursk Operation Simulation and Validation Exercise) study. The data are two-sided, time phased (daily), and highly detailed. They cover 15 days of the Battle of Kursk.

A total of 39 diverse models are fit to the Battle of Kursk data using different approaches. These approaches include applying the methodologies of previous studies, using robust LTS (least trimmed squares) regression for estimation purposes, including air sortie data of the battle, considering the battle in separate phases, using different weights to aggregate the data, fitting basic Lanchester equations, fitting Morse-Kimball equations, and applying parameters found in previous studies.

The findings from this research include:

- It is observed that the original Lanchester equations do not fit to the Battle of Kursk data, and therefore may not be appropriate for modeling the combat. Of the three ill-fitting Lanchester equations, the best fit is obtained by applying the linear law, which is used for modeling ancient warfare or area fire.
- The parameters derived from Bracken and Fricker's Ardennes studies do not apply to the Battle of Kursk data. This implies that there are no unique parameters that apply to all battles.
- Throughout the study, the *a* parameter is generally greater than the *b* parameter. This implies that individually German soldiers were more lethal than Soviet soldiers.
- The best fit to the data is observed when a robust LTS regression model is applied. The best fit occurred with no attacker/defender advantage.
- In the battle of Kursk, except for the first and eighth days, it was advantageous to be the attacker.
- The different approaches give very different estimates on the best fitting Lanchester parameters (p, q, a, b); see Table 1 for a sample of the range of

parameters found. A closer investigation reveals that the surface of the sum squared residuals (SSR), from the differences between the daily estimated and real losses over the battle, is very flat. This explains why such diverse answers are in the literature. Figure 1 shows a contour filled plot of SSR values for Battle of Kursk data with p values varied between –0.5 and 10.0 and q values varied between –1.0 and 4.0, and a and b values determined to minimize SSR given p and q. Different researchers, using different methods, all came up with completely different answers because the surface around the models' fits is very flat. Therefore, small changes in handling the data and the application of different estimation methodologies results in dramatically different parameter estimates. . Thus, there is not enough data from the battle of Kursk, and Ardennes too, to differentiate between a wide range of Lanchester models. Unfortunately, none of the basic Lanchester models, liner, square, and logarithmic, provide a good fit.

Name of the model	***a***	***b***	***p***	***q***
Bracken Ardennes Model 1	8.0E-9	1.0E-8	1.0	1.0
Bracken Ardennes Model 3	8.0E-9	1.0E-8	1.3	0.7
Fricker Ardennes Com.Manpwr.w/o sortie	4.7E-27	3.1E-26	0.0	5.0
Clemens Kursk Linear Regression	6.92E-49	6.94E-48	5.3157	3.6339
Clemens Kursk Newt.-Raphson iteration	3.73E-6	5.91E-6	0.0	1.6178
Kursk Robust LTS Regression	2.27E-40	1.84E-41	6.0843	1.7312

Table 1. Results for the 6 of the 39 models explored in the study. Notice the wide range of parameter estimates.

- Combat models cannot provide clear-cut results to a military analyst. One cannot determine the outcome of a battle precisely by using combat models. Together with their use to gain insight about the battles and campaigns that happened in the past, combat models help to make better decisions by enabling the decision-maker to compare different alternatives using various combat modeling techniques.

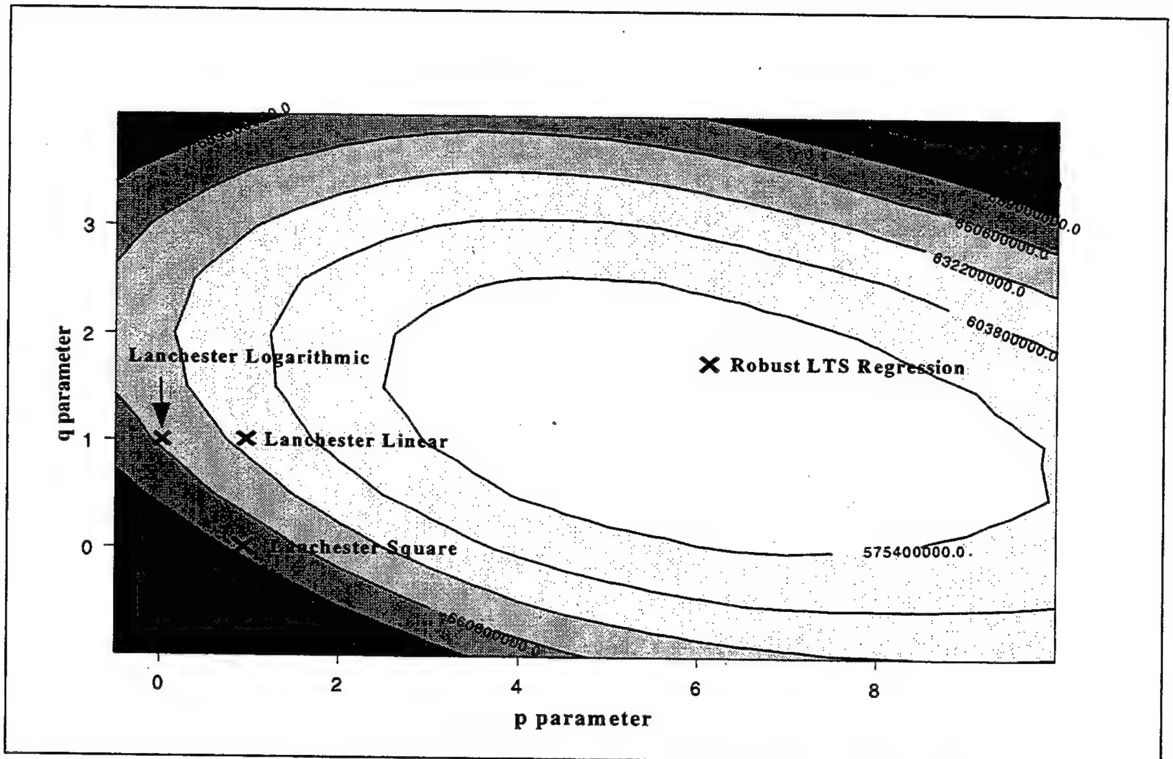

Figure 1. Contour filled plot of SSR values for Battle of Kursk data with no attacker/defender advantage considered. A wide range of diverse generalized Lanchester models give about the same fit.

LIST OF SYMBOLS, ACRONYMS AND ABBREVIATIONS

- SP : Self Propelled
- KDB : Kursk Database
- OH : On hand
- HQ : Headquarters
- KIA : Killed In Action
- WIA : Wounded In Action
- CMIA : Captued/Missing In Action
- DNBI : Disease/Nonbattle Injuries

I. INTRODUCTION

War is a conflict between nations or states carried on by force of considerable duration and magnitude, by land, sea, or air for obtaining and establishing the superiority and dominion of one over the other for some cause. It is defined more concisely as the state of usually open and declared armed hostile conflict between states or nations [Ref.1]. When these conflicts reach global proportions, they are known as world wars. Among the causes of war are ideological, political, racial, economic, and religious conflicts. According to Karl von Clausewitz, war is a "continuation of political intercourse by other means" and often occurs after means of compromise and mediation have failed.

Throughout history, war has been a topic of analysis for scientists and researchers, especially following World War II. In the shadow of a possible outbreak of nuclear war between the United States and Russia, more research has been done on the subject of war than ever before.

A. COMBAT MODELING

This study, instead of analyzing the concept of war at large, will analyze and focus on a smaller part of war, which we usually name *combat*, *battle* or *campaign*. Even though these terms are generally used as if they had the same meaning, the word *combat* is used for defining active, armed fighting between two enemy forces, while the word *battle* is used for defining a hostile encounter or engagement between opposing military forces. The word *campaign* is used for defining military operations for a specific objective, and defines a connected series of military operations aimed at accomplishing a

specific operational and strategic objective. A *campaign* forms a separate and distinct phase of war, and it is this "small" part of war that is the concentration of this study.

Throughout history, combat has been an important topic of analysis, just like war itself. Scientists, researchers, and the military have tried to understand and estimate beforehand the nature of combat in order to formulate some theory about its dynamics and most importantly, its outcome. Researchers who studied combat modeling and attrition were aware of the influence their studies could have on the outcome of a battle. A natural consequence of these studies was the emergence of combat models in the beginning of the early 20th century.

Attrition is a reduction or decrease in number, size, or strength of a force and is at the core of every general discussion of warfare. The term attrition defines a wearing down or weakening of resistance, especially as a result of constant harassment, abuse, or attack.

Soviets argue that Osipov [Ref.2] was the first to study and discover the equations most often used when modeling attrition in combat. The equations are widely known as, "Lanchester's equations." Regardless of claims of prior or parallel discovery, Lanchester's equations for attrition provided the origin for modeling attrition in the United States and around the world.

Frederick William Lanchester (b.1868, London, England; d.1946, Birmingham, Warwickshire) was an English automobile and aeronautics pioneer who built the first British automobile in 1896. Lanchester's interest in aeronautics was first expressed in a paper he wrote in 1897, a work ahead of its time discussing the principles of heavier-than-air flight. Between 1907-1908, he published a two-volume work embodying

distinctly advanced aerodynamic ideas. As a member of the Advisory Committee on Aeronautics in 1909 and, later, as a consultant to the Daimler Motor Company, Ltd., Lanchester also contributed to the development of the field of operations research. [Ref.3].

Lanchester proposed that attrition could be mathematically modeled, and introduced his equations as a means of investigating the future impact that the recently invented airplane might have on the nature of warfare [Ref.4]. Thus, at the beginning of World War II, Lanchester equations and other differential equations of a similar nature were known to some of the scientists who later became active in operations research [Ref.5].

Today, with the advent of computers, Lanchester-based models of warfare are widely used in the decision making process for research, development, acquisition of weapons systems, force mix decisions, and for aiding in the development of operational plans.

B. LANCHESTER EQUATIONS

As described in Fricker [Ref.6], the basic generalized Lanchester Equations are of the form:

$$\dot{B}(t) = aR(t)^p B(t)^q \qquad (1)$$

$$\dot{R}(t) = bB(t)^p R(t)^q \qquad (2)$$

where $B(t)$ and $R(t)$ are the strengths of blue and red forces at time t, $\dot{B}(t)$ and $\dot{R}(t)$ are the rates at which blue and red force levels are changing at time t, a and b are attrition parameters, p is the exponent parameter of the attacking force, and q is the exponent parameter of the defending force. The model begins with initial force sizes, $B(0)$ and

$R(0)$, that, when solved numerically, are incrementally decreased according to the relationship $B(t+\Delta t) = B(t) - \Delta t\,\dot{B}(t)$ and $R(t+\Delta t) = R(t) - \Delta t\,\dot{R}(t)$. In an equally matched battle, where the ratio of the forces stays constant over time, $B(t)/R(t) = \dot{B}(t)/\dot{R}(t)$, for all t. This is equivalent to the condition that $bB(t)^{p-q+1} = aR(t)^{p-q+1}$ for some p and q, and all t.

Two versions of the Lanchester equations are of particular interest. When $p = q = 1$ (or, more generally, when $p-q = 0$) force ratios remain equal if $aR(0) = bB(0)$, and hence this condition is called, *Lanchester's linear law*. The interpretation of Lanchester's linear law is that a battle governed by this model is characterized as a collection of small engagements, and was proposed by Lanchester [Ref.4] as a model for ancient warfare. The equation is also considered a good model for area fire weapons, such as artillery [Ref.7].

Lanchester contrasted the Linear Law with the condition $p = 1$, $q = 0$ (or, more generally, $p-q = 1$), which is called *Lanchester's square law*, where the force ratios remain equal when $aR(0)^2 = bB(0)^2$. He theorized that the square law applies to modern warfare, in which both sides are able to aim their fire. His model suggests that in modern warfare, combatants should concentrate their forces.

A third version with $p = 0$, $q = 1$ (or, more generally, $q-p = 1$) is called *Lanchester's logarithmic law*.

C. THESIS OUTLINE

This thesis consists of five chapters. This first chapter introduces the general concept of combat modeling and the widely used Lanchester Equations. The second chapter reviews previous studies on combat modeling. The Battle of Kursk data is also

introduced in this chapter, and the study methodology for the thesis is explained. The third chapter briefly covers the history of Battle of Kursk, and explores and analyzes the battle's data in depth to gain insights before attempting to fit models to it. Additional details about the personnel and weapon systems data is given in Appendix A.

The primary objective of Chapter Four is to find the best model that fits the Battle of Kursk data. To accomplish this objective, the methods of previous studies are applied, and then all new exploratory models are implemented. The results derived from the regression analysis methods are briefly evaluated in this chapter. The fifth chapter interprets the results that are derived from the regression analysis methods in Chapter Four. Chapter Five contains the final conclusions and recommendations implicated by results, and also mentions future areas of study on combat modeling.

THIS PAGE IS LEFT INTENTIONALLY BLANK

II. PREVIOUS STUDIES ON COMBAT MODELING

A. PREVIOUS STUDIES WITH LANCHESTER EQUATIONS

Past empirical validation studies of Lanchester Equations include the work of Bracken [Ref.8] on the Ardennes campaign of World War II, Fricker [Ref.6], also on the Ardennes campaign, Clemens [Ref.9] on the Battle of Kursk of World War II, and Hartley and Helmbold [Ref.10] on the Inchon-Seoul campaign of the Korean War. These works are among the few quantitative studies that use daily force size data for real battles.

1. Bracken's study

Bracken formulates four different models [Ref.8] for the Ardennes campaign, which are variations of basic Lanchester equations, and estimates their parameters for the first ten days of the of the Ardennes campaign of World War II (December 15, 1944 through January 16, 1945).

Bracken's models are homogeneous. Tanks, armored personnel carriers, artillery, and manpower are aggregated with weights representing the relative effectiveness of the weapon systems. This type of aggregation yields a single measure of strength for each of the Allied and German forces. This method is used to measure combat power and to calculate losses. His models treat combat forces and the total forces (i.e., both support forces and the combat forces) in the campaign separately.

Equations II.A.1.(3), II.A.1.(4) show the Lanchester equations used by Bracken, which are modified to include the tactical parameter d for Bracken's Model 1 and Model 2. The parameter d is a multiplier of attrition due to being either in a defensive or offensive posture in the battle. If $d < 1$, then the defender has fewer casualties (i.e., there is a defender advantage). If $d > 1$ then the defender has more casualties (i.e., there is an

attacker advantage). If d=1 then there is no attacker or defender advantage. Using the tactical parameter *d* requires knowing which side is the defender and which side is the attacker.

$$\dot{B} = (d \; or 1/d) \; a \; R^p B^q \tag{3}$$

$$\dot{R} = (1/d \; or \; d) \; b \; B^p R^q \tag{4}$$

In Model 1, forces are composed of tanks, APCs, artillery, and combat manpower; where combat manpower is made up of infantry, armor, and artillery personnel. Manpower casualties are killed and wounded. Forces are tanks, APCs, artillery, and combat manpower, which are weighted by 20, 5, 40, and 1, respectively. That is, *Blue Forces (combat power) = (20* x *number of tanks) + (5* x *number of APCs) + (40* x *number of artillery) + (1* x *number of combat manpower).* Bracken [Ref.8] states in his study that, "The weights given above are consistent with those of studies and models of the U.S. Army Concepts Analysis Agency. Virtually all theater-level dynamic combat simulation models incorporate similar weights, either as inputs or as decision parameters computed as the simulations progress."

In Model 2, forces include all personnel in the campaign, including all types of logistics and support personnel. Casualties are personnel who are killed, wounded, captured or missing in action, and who have disease and nonbattle injuries. It is noteworthy here to mention that in the Ardennes campaign, the Allies had a smaller portion of their forces in combat units and a larger portion of their forces in logistics and support units than the Germans.

In estimating the parameters of Model 1, Bracken found that individual German effectiveness, as measured by the attrition parameter *a*, is less than Allied effectiveness *b*;

these parameters are for combat forces only. This distinction is a natural result of the German combat forces having less support, and therefore not being as effective as Allied combat forces individually. In Model 2 where all personnel are included, individual effectiveness is determined to be similar for both the Allied forces and the Germans.

In Model 3, the components used are the same as in Model 1, but the parameter *d* is not estimated. Just like Model 3, Model 4 does not have a tactical parameter. Model 4, like Model 2, addresses total forces rather than combat forces. For a summary of Bracken's models, see Table 2.

	COMBAT MANPOWER	SUPPORT MANPOWER	PARAMETER *d*
MODEL1	X		X
MODEL2	X	X	X
MODEL3	X		
MODEL4	X	X	

Table 2. Bracken's models summarized. Model 1 and Model 3 use combat manpower only; Model 2 and Model 4 use total manpower. Combat manpower is made up of infantry, armor, and artillery personnel; support manpower is made up of all types of logistics and support personnel. Model 1 and Model 2 have defensive parameter *d*; Model 3 and Model 4 do not have *d*.

Bracken's main conclusions are:

- Lanchester linear model best fits the Ardennes campaign data in all four cases.
- When combat forces are considered, Allied individual effectiveness is greater than German individual effectiveness. When total forces are considered, individual effectiveness is the same for both sides.
- There is an attacker advantage.

The second result indicates that the two sides have essentially the same individual capabilities but are organized differently. The Allies preferred to have more manpower

in the support forces, which in turn yielded greater individual capabilities in the combat forces. The overall superiority of the Allied forces in the campaign led to the Allied attrition being a smaller percentage of their forces. Table 3 shows Bracken's best fitting parameters for the Ardennes campaign.

Name of the model	*a*	*b*	*p*	*q*	*d*
Bracken Model 1	8.0E-9	1.0E-8	1.0	1.0	1.25
Bracken Model 2	8.0E-9	8.0E-9	0.8	1.2	1.25
Bracken Model 3	8.0E-9	1.0E-8	1.3	0.7	-
Bracken Model 4	8.0E-9	8.0E-9	1.2	0.8	-

Table 3. Bracken's parameters found in his study for Ardennes campaign data.

2. Fricker's study

Fricker's paper [Ref.6] revisits Bracken's modeling of the Ardennes campaign of World War II [Ref.8] and uses the Lanchester equations. This is different than Bracken's study in several ways. Fricker's study:

- Uses linear regression to fit the model parameters.
- Uses the total body of data from the entire campaign, while Bracken used only the first 10 days of the data from the Ardennes Campaign.
- Also includes air sortie data.

In contrast to Bracken, Fricker shows that the Lanchester linear and square laws do not fit the data. He concludes by showing that a new form of the Lanchester equations—with a physical interpretation—fits best. Fricker states that the attrition

parameter used in the Lanchester logarithmic model represents the opponent's probability of killing a soldier, and that this probability of kill is constant for a certain range of the opponent's force sizes. It follows that one side's losses are more a function of own forces rather than a result of the opponent's forces, and Fricker gives the Gulf War as support for this theory. That is, Iraqi casualties were more a function of the number of Iraqi forces than of the number of Allied forces. Table 4 shows the best fitting parameters for the Ardennes campaign according to Fricker's study.

Name of the model	***a***	***b***	***p***	***q***	***d***
Combat manpower w/o sortie	4.7E-27	3.1E-26	0.0	5.0	0.8093
Total manpower w/o sortie	1.7E-16	8.0E-16	0.0	3.2	0.824
Combat manpower With sortie	2.7E-24	1.6E-23	0.0	4.6	0.7971
Total manpower with sortie	1.3E-15	5.6E-15	0.0	3.0	0.8197

Table 4. Fricker's parameters from his study of the Ardennes campaign data. The estimated *d* parameter indicates a defender advantage. The *d* parameter used in Fricker's study is the inverse of the *d* parameter defined in Bracken's study.

3. Clemens' study

Clemens' analysis [Ref.9] examines the validity of the Lanchester Models as they are applied to modern warfare. The models in his study are based upon basic Lanchester Equations. The analysis is an extension of Bracken's [Ref.8] and Fricker's [Ref.6] analyses of the Ardennes Campaign, and applies the Lanchester models to the Battle of Kursk data.

Clemens uses two estimation techniques, linear regression and Newton-Raphson iteration. The analysis also explores the presented model in matrix form, and compares the matrix solution to the scalar solution. In his study he concludes that:

- Neither the Lanchester linear nor the Lanchester square model fits the data.
- The Lanchester logarithmic model in both scalar and matrix form fits better than the Lanchester linear and square models.
- Lanchester Equations do not give the best fit for the data.
- The analysis can be extended by:
    - Taking into account the change in offensive/defensive roles.
    - Adding data from air sorties.
    - Applying the Lanchester Equations in a homogeneous weapon scenario.
    - Building a whole new model without regard to the Lanchester formulations.

Table 5 shows the best fitting parameters Clemens found for the Battle of Kursk data in his study.

Name of the model	***a***	***b***	***p***	***q***	***d***
Clemens Linear Regression	6.92E-49	6.94E-48	5.3157	3.6339	-
Clemens Newton-Raphson	3.73E-6	5.91E-6	0.0	1.6178	-

Table 5. Clemens' parameters found in his study for the Battle of Kursk Data.

4. Hartley and Helmbold's study

Hartley and Helmbold's study [Ref.10] focuses on validating the homogenous Lanchester square law by using historical combat data. Since validating a model means testing it in a real life context, Hartley and Helmbold test Lanchester's square law using the data from the Inchon-Seoul campaign of the Korean War.

Hartley and Helmbold use three analysis techniques to examine the data; linear regression, the Akaike Information Criterion (AIC), and Bozdogan's consistent AIC (CAIC). The results of the study are:

- The data do not fit a constant coefficient Lanchester square law.
- The data better fit a set of three separate battles (one distinct battle every six or seven days). However, the data fit a set of three constant casualty-model battles just as well.
- Lanchester square law is not a proven attrition algorithm for warfare, but neither can it be completely discounted.
- More real combat data are needed to validate any proposed attrition law such as the Lanchester square law.

5. A summary of previous findings

Fricker's and Bracken's studies are significant in that they reach different conclusions using the same data. When both studies are compared, Fricker's approach and methodology makes more sense because he did not constrain himself to certain ranges of parameters, as Bracken did.

Bracken's approach is strong in the sense that his approach optimizes the nonlinear regression equation in the defined area. Fricker finds the parameters that give

the minimum sum of squared residuals (SSR), using the logarithmically transformed Lanchester equations. Using logarithmic transformation does not necessarily guarantee the best fit when the parameters found by this approach are directly applied to the Lanchester equations. However, minimizing the SSR value was Bracken's criteria and the parameters found via logarithmic transformation in Fricker always resulted in smaller sums of square errors for the untransformed Lanchester equations than those found by Bracken.

In general, the results of all four studies show no overwhelming evidence of Lanchester fit. Among the three Lanchester equations, the logarithmic law gives the best fit.

B. THE DATA AND STUDY METHODOLOGY

1. The data

Complete combat data on both sides fighting against each other is very sparse. Consequently, validation of Lanchester and other combat models has been very difficult, and the most accessible battle data contains only starting sizes and casualties, sometimes only for one side. Furthermore, the definition of casualties varies (e.g., killed, killed plus wounded, killed plus wounded plus missing, killed plus wounded plus missing plus disease/nonbattle injuries), making data analysis difficult. Obtaining order-of-battle data and equipment damage reports requires extensive historical research. Recently, more data has become available and improved database management and computing power has helped in such data gathering efforts.

A detailed database of the Battle of Kursk of World War II, the largest tank battle in history, was recently developed. The data were collected from military archives in

Germany and Russia by the Dupuy Institiute (TDI), and are reformatted into a computerized data base, designated as the Kursk Data Base (KDB). The KDB was recently documented in the KOSAVE (Kursk Operation Simulation and Validation Exercise) study. [Ref.12]. The data are two-sided, time phased (daily), and highly detailed. They cover 15 days of the Battle of Kursk.

2. Study methodology

This thesis fits Lanchester equations and other functional forms to the newly released Battle of Kursk data. The two main areas of interest are the quality of the fits and the insights provided by the equations. Different fits are compared and contrasted to the previous research results mentioned above.

The methodology used in this thesis research consists of the following steps and research questions:

- Arranging and setting up of the data at hand so that it is useful for regression and statistical purposes.
- Conducting a thorough analysis and interpretation of the data.
- Identifying components needed for the model.
- Applying Bracken's and Fricker's methodology to the Kursk data.
- Applying various forms of Lanchester Equations to the data. How well do Lanchester Equations fit the Battle of Kursk Data?
    - Does the Linear Law fit the Battle of Kursk data?
    - Does the Square Law fit the Battle of Kursk data?
    - Does the Logarithmic Law fit the Battle of Kursk data?

- Do possible combinations of these three laws fit the Battle of Kursk data?

- Applying other general curve fittings and functional forms to the data.
- Do any of the other possible general curve fits or functional forms fit the Battle of Kursk data?
  - Do any of the functional forms need the defender/attacker coefficient?
  - What effect does changing weapon weights have on fitting the models to the data?
- Using a least squares grid search to get a better understanding of the relationship between various Lanchester formulation and the empirical data.
- Comparing and contrasting different methodologies and the two battles.
- Analyzing the results and conclusions of all the models.

This thesis extends the previous studies of Bracken, Fricker, Clemens, and Hartley and Helmbold in the following ways:

- Methodologies of previous studies are applied to Battle of Kursk data.
- A different regression technique, i.e., robust LTS regression, is used.
- Air sortie data is included.
- The change in offensive/defensive roles is taken into account.
- The battle is considered in different phases and different change points are used for fitting the model.
- Different weights are used for aggregating the data.

- Lanchester Equations, Morse-Kimball equations and force ratio models are fit to Battle of Kursk data
- Parameters found for different battles are used to fit Battle of Kursk data and the resulting parameters are compared and contrasted. By this comparison, the issue of whether or not the parameters of one battle can be used for another battle is discussed.

THIS PAGE IS LEFT INTENTIONALLY BLANK

III. HISTORY AND THE DATA OF THE BATTLE OF KURSK

A. SHORT HISTORY OF THE BATTLE OF KURSK

1. City of Kursk

The city of Kursk is the administrative center of Kursk Province in western Russia. It lies along the upper Seym River, about 280 miles south of Moscow. First mentioned in documents from 1032, Kursk is one of the oldest cities in Russia (Figure 2). Completely destroyed by the Tatars in 1240, Kursk was not rebuilt until 1586, when it became a military outpost protecting the advancing Russian colonization from Tatar attack. The town, however, lost much of its importance at the beginning of the 18th century when the Russian border was moved farther south. In World War II, fierce fighting took place around Kursk and the city was severely damaged (Figure 3). In July, 1943, the Battle of Kursk, the largest tank battle in World War II, took place around the city of Kursk and ended in the defeat of the Germans. [Ref.13].

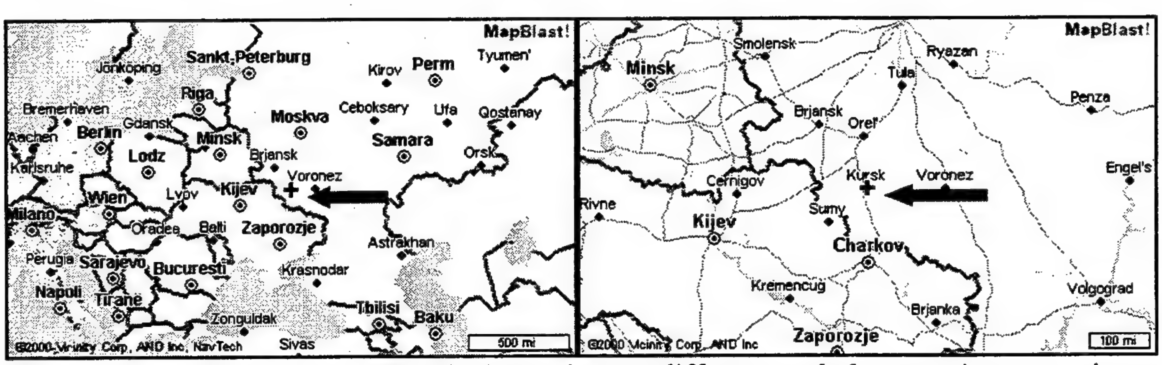

Figure 2. Location of city of Kursk shown in two different scaled maps. Arrows point to the plus signs showing the city's exact location. [Ref.14][Ref.15].

2. The history of the Battle of Kursk

During World War II, following the German defeat in Stalingrad, the military situation in the Eastern Front was much different than it was the year before. After

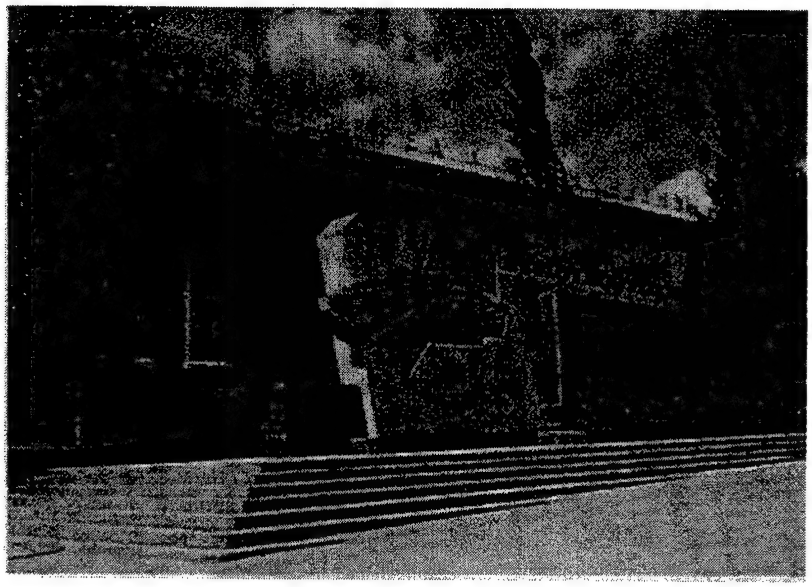

Figure 3. *Monument to World War II dead* in the city of Kursk, Russia [Ref.13]

Stalingrad, the Russians knew they were going to win the war, and the Germans strongly suspected they might lose the war. The war was in mid-course—the outcome might be predictable, but the Germans were far from beaten. The Germans were still mighty, powerful and dangerous. In the spring of 1943, the Eastern Front was dominated by a salient located to the north of city of Kharkov, to the south of city of Orel, and centered on the city of Kursk. The Kursk salient was 250 miles wide and 70 miles deep. The German plan was a two-front attack on the Kursk salient in a classic pincer operation.

Operation Citadel was launched on July 5, 1943. On July 2, 1943, Adollf Hitler said, "This attack is of decisive importance. It must succeed, and it must do so rapidly and convincingly. It must secure for us the initiative for this spring and summer. The victory of Kursk must be a blazing torch to the world." [Ref.16: p.103].

With the objectives of destroying Soviet forces and eliminating the salient by linking up the area around the city of Kursk, General Model's 9th Army attacked from the north, while General Hoth's 4th Panzer Army attacked from south of the salient. The

Soviets had enough time to prepare heavily fortified defense lines because of frequent German planning delays, and this advantage was a major setback for the Germans that contributed to their defeat in this battle. (See the map in Figure 4).

General Model's 9th Army's attack from the north gained approximately 6 miles of ground into the enemy lines before being stopped on Day 4. Following Day 4, the German attack on the northern front was stalled. General Hoth's 4th Panzer Army's attack from the south was more successful. Following an initial gain of a few miles in the first two days of the battle, the 4th Panzer Army caused great damage and alarm among the Soviets. Despite their heavy losses, the Soviets were able to restrict German progress to a mere 25 miles by July 12. A German breakthrough attempt on July 12 resulted in the greatest single armored engagement in history near the town of Prokhorovka, when Germans ran into the advancing 5th Guards Tank Army, which was the Soviet theater reserve (i.e., the biggest Soviet reserve force in the battlefield at the time).

As night closed over Prokhorovka, the greatest armored battle in history had fought itself out. The field was strewn with more than 300 German tanks, including 70 of the huge Tigers, 88 SP guns and 300 trucks. Rotmisrov's 5th Guards Tank Army had suffered a 50 percent loss of his 850 tanks and SP guns. The dazed Germans described the day as the *Bluthmähle von Belgorod* (the bloodbath at Belgorod). Unable to gain a decisive victory, and stopped by Soviet reinforcements and counterattacks, the Germans drew back into defensive postures after this battle. [Ref.16][Ref.17].

While the number of German tank losses cited in historical sources is around 300, this is different than the number of German tank losses given in the KDB, which is 98.

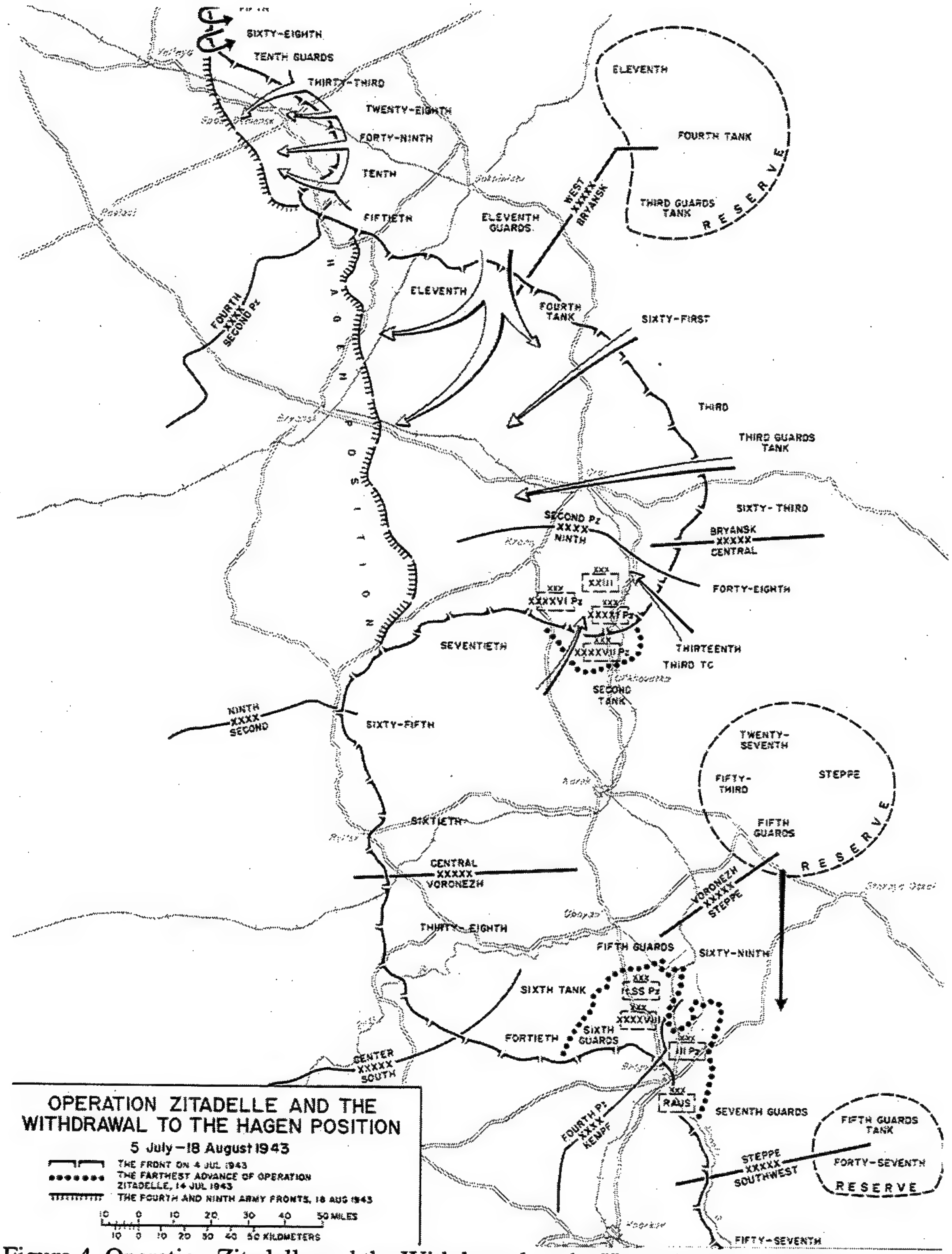

Figure 4. Operation Zitadelle and the Withdrawal to the Hagen Position. [Ref.17 : p.134]

Hitler's orders to cancel Citadel on July 13, 1943 came as a shock to the German field commanders, and consequently, further attacks were limited in scope. The Soviets began their counterattack on the southern front on July 12, and regained all ground lost in the theater by July 23, 1943.

Soviet military historians named the Battle of Kursk as "the Nazi Waterloo". Following this Battle, the military situation in the Eastern Front got worse for the Germans. Detailed information about the military aspects of the Battle of Kursk can be found in [Ref.16] and [Ref.17].

B. BATTLE OF KURSK DATA

The data presented in KOSAVE (Kursk Operation Simulation and Validation Exercise) [Ref.12] consists of mainly 6 parts:

- Units and combat posture status.
- Personnel status and casualties.
- Army weapons status and losses.
- Ammunition status.
- Aircraft sortie status.
- Geographic unit positions and progress.

1. Limitations and timeframe of the Kursk database

The KDB in the KOSAVE report includes quantified data only on the southern front of the Kursk Battle. Results are not expressed in terms of specific weapon types, and weapons are aggregated into categories or classes for tractability. Human factors like fatigue, morale, caution, aggressiveness, regulating the pace and intensity of battle are not

quantified. The data given is for 15 days of battle, from 4 July 1943 through 18 July 1943. [Ref.12].

2. Assumptions made for Kursk database

The KDB accurately represents the status and structure of forces in the southern front of the actual World War II Battle of Kursk. The personnel casualty and system kill criteria used to categorize KDB casualty and weapon losses are sufficiently consistent with each other to allow meaningful reporting and comparisons between combatants. The use of interpolation techniques for gathering data between inconsistent reports in historical records create a complete set of daily report records in the KDB is reasonable. [Ref.12].

3. Phases of the battle

With the start of the battle on July 4, 1943, German forces encountered heavy losses as they fell upon the fortified Soviet positions. The Germans were able to advance only 10 miles south and 30 miles north into the salient before the offensive stalled. The Soviets mounted their counterattack on July 12, and by July 18, had decisively won the battle. Soviets retained the initiative and used it to dominate the Eastern Front until the end of the war. In summary, the days of attack and defense are:

- July 4 - July 11 = day 1- day 8 of the battle = Germans attack
- July 12 – July 18 = day 9- day15 of the battle = Soviets attack

Throughout the thesis, the Soviet forces are known as Blue forces German forces as Red forces.

C. METHODOLOGY USED FOR GATHERING DATA

The critical data is extracted from the KOSAVE [Ref.13] report, and aggregated depending on the model used in the study. This process was the most difficult and time-consuming process of the study. Extracting only that information needed from an immense database demands great attention to detail and methodology. The methodology of how the data is gathered for modeling purposes is explained in detail in the section concerning that specific model. The general outlines, which do not change for every model, are explained in III.C.1 and III.C.2. All data used in this study are for combat units represented in the KDB. Support units, such as bridging and logistics units, are excluded [Ref.12].

1. Manpower data

Throughout the study, combat manpower is used for modeling the combat. The manpower presented as "On Hand" (OH) in the KOSAVE [Ref.12] report is summed up, including the headquarters units, and is assumed to be the number of combat forces. Thus the number of combat forces is assumed to represent all the combat forces available on hand. Combat manpower losses are killed, wounded, captured/missing in action, and disease and nonbattle injuries.

Table 6 shows the combat manpower data for the Soviet forces and Table 7 shows the combat manpower data for the German forces.

2. Weapon systems data

Throughout this study, the total number of weapon systems is used for modeling purposes. The weapon systems presented as OH in the KOSAVE [Ref.12] report are summed up, including the weapon systems of headquarter units, and assumed to be the

total number weapon systems. The total number of weapon systems is assumed to represent all the weapon systems available on hand, including the weapon systems of headquarter units.

Day	OH Manpower	KIA	WIA	CMIA	DNBI	Total
1	510252	30	73	11	16	130
2	507698	1616	3548	3281	82	8527
3	498884	1911	3861	3553	98	9423
4	489175	2160	4949	3230	92	10431
5	481947	2069	4767	2585	126	9547
6	470762	2613	6451	2561	211	11836
7	460808	2326	5189	3209	46	10770
8	453126	1792	4488	1417	57	7754
9	433813	4417	11450	3496	59	19422
10	423351	2205	6709	1556	52	10522
11	415254	2153	5315	1206	49	8723
12	419374	822	2641	575	38	4076
13	416666	619	1928	358	35	2940
14	415461	225	846	111	35	1217
15	413298	881	2198	151	30	3260
TOTAL		**25839**	**64413**	**27300**	**1026**	**118578**

Table 6. Soviet combat manpower data. KIA denotes killed in action, WIA denotes wounded in action, CMIA denotes captured/missing in action and DNBI denotes disease and nonbattle injuries.

Day	OH Manpower	KIA	WIA	CMIA	DNBI	Total
1	307365	129	516	12	143	800
2	301341	960	4817	272	143	6192
3	297205	565	3375	217	145	4302
4	293960	475	2726	70	143	3414
5	306659	503	2202	75	162	2942
6	303879	492	2201	100	160	2953
7	302014	304	1532	43	161	2040
8	300050	345	1893	68	169	2475
9	298710	420	1944	86	162	2612
10	299369	327	1533	33	158	2051
11	297395	338	1584	63	155	2140
12	296237	220	912	31	159	1322
13	296426	214	920	27	189	1350
14	296350	138	622	26	163	949
15	295750	161	707	19	167	1054
TOTAL		**5591**	**27484**	**1142**	**2379**	**36596**

Table 7. German combat manpower data. KIA denotes killed in action, WIA denotes wounded in action, CMIA denotes captured/missing in action and DNBI denotes disease and nonbattle injuries.

Weapon losses are destroyed/abandoned and damaged. In the example presented, considering a damaged weapon system as a loss is logical, because a damaged weapon system is considered to be a "temporary" loss and in a non-operational status. Therefore, damaged weapons are also included when calculating the losses of both sides. A damaged weapon system is treated as only a "temporary" loss, but the period of non-operational status can be long. Also, a damaged operational system will often function only with degraded effectiveness and efficiency. [Ref.12: p.5-13].

Likewise, since a damaged weapon is damaged "in action" and is left out of the battle indefinitely due to its non-operational status until repaired, it will be considered as a loss. "Damaged" denotes number of items damaged in action. [Ref.12: p. H-1].

In order to prevent possible confusion for future analysts, the methodology used while gathering data for the classification of weapon systems will be offered. The type of each weapon system is listed below. Table 5-1 [Ref.12: p.5-3], Table 5-2 [Ref.12: p.5-4], and the tables from the weapons lists in the Appendices of the KOSAVE II Study Report [Ref.12] are used for purposes of classification. The results are as follows.

a. Classification of Soviet weapon systems

(1) Tanks used in the study:

- KV-1, KV-2
- M-3, MK-2/3, MK-4
- T-34, T-60, T-70

(2) APCs used in the study:

- BA-64, BA-10
- Armtpt

- Bren

(3) Artillery used in the study report:

- SU-122
- 122mm Gun
- 122mm How
- 152mm Gun
- SU-152
- 203mm How

b. *Classification of German weapon systems*

(1) Tanks used in the study:

- PzIII(all types), PzIV(all types), PzV(all types), PzVI(all types)
- T-34(Soviet), PzIIIspt

(2) APCs used in the study:

- AC4-6w, AC8w
- LHT, MHT, LHTspt, MHTspt
- Acspt, AC8w 75mm
- MHT75mmIG
- Pz I, Pz II
- MHT Flame

(3) Artillery used in the study:

- 75mm lt IG
- 105mm Gun, 150mm Gun

- 87.6mm How, 105mm How(towed and SP),
- 150mm How, 152mm How, 155mm How, 210mm How
- Wespe (is a SP Artillery Gun with 105mm Light Field Howitzer)
- Hummel (is a 150mm SP Gun)

Using the methodology explained above, the necessary data is gathered from the KOSAVE [Ref.12] report. The data for the tank, APC and the artillery weapon systems for the Soviet forces are given in Table 8, Table 9 and Table 10, consecutively. The data for the tank, APC and the artillery weapon systems for the German forces are given in Table 11, Table 12 and Table 13, consecutively.

Day	OH Tanks	Damaged	Dst+Abnd	Total Loss
1	2500	0	0	0
2	2396	19	86	105
3	2367	69	48	117
4	2064	120	139	259
5	1754	100	215	315
6	1495	149	140	289
7	1406	77	80	157
8	1351	51	84	135
9	977	210	204	414
10	978	58	59	117
11	907	57	61	118
12	883	45	51	96
13	985	9	18	27
14	978	16	26	42
15	948	58	27	85
TOTAL		**1038**	**1238**	**2276**

Table 8. Soviet tank data. Dst+Abnd denotes destroyed and abandoned tanks. OH denotes the on hand amount.

Day	OH APC	Damaged	Dst+Abnd	Total Loss
1	511	0	0	0
2	507	0	4	4
3	501	0	6	6
4	490	2	9	11
5	477	1	12	13
6	458	12	7	19
7	463	0	3	3
8	462	2	2	4
9	432	15	15	30
10	424	2	6	8
11	418	2	6	8
12	417	0	1	1
13	417	0	0	0
14	417	2	0	2
15	409	6	2	8
TOTAL		**44**	**73**	**117**

Table 9. Soviet APC data. Dst+Abnd denotes destroyed and abandoned APCs. OH denotes the on hand amount.

Day	OH Artillery	Damaged	Dst+Abnd	Total Loss
1	718	0	0	0
2	705	2	11	13
3	676	2	28	30
4	661	7	8	15
5	648	2	12	14
6	640	0	9	9
7	629	1	12	13
8	628	1	6	7
9	613	2	14	16
10	606	3	7	10
11	603	0	5	5
12	601	2	3	5
13	600	0	3	3
14	602	0	0	0
15	591	0	4	4
TOTAL		**22**	**122**	**144**

Table 10. Soviet artillery data. Dst+Abnd denotes destroyed and abandoned artillery. OH denotes the on hand amount.

Day	OH Tanks	Damaged	Dst+Abnd	Total Loss
1	1178	2	2	4
2	986	175	23	198
3	749	216	32	248
4	673	107	14	121
5	596	92	16	108
6	490	107	32	139
7	548	32	4	36
8	563	48	15	63
9	500	89	9	98
10	495	50	7	57
11	480	32	14	46
12	426	70	9	79
13	495	15	8	23
14	557	6	1	7
15	588	6	0	6
TOTAL		**1047**	**186**	**1233**

Table 11. German tank data. Dst+Abnd denotes destroyed and abandoned tanks. OH denotes the on hand amount.

Day	OH APC	Damaged	Dst+Abnd	Total Loss
1	1170	0	0	0
2	1142	23	6	29
3	1128	13	1	14
4	1101	16	11	27
5	1085	14	2	16
6	1073	12	2	14
7	1114	33	9	42
8	1104	12	4	16
9	1099	3	9	12
10	1096	3	1	4
11	1093	5	1	6
12	1089	4	1	5
13	1092	0	1	1
14	1095	0	1	1
15	1098	4	1	5
TOTAL		**142**	**50**	**192**

Table 12. German APC data. Dst+Abnd denotes destroyed and abandoned APCs. OH denotes the on hand amount.

Day	OH Artillery	Damaged	Dst+Abnd	Total Loss
1	1189	1	0	1
2	1166	10	14	24
3	1161	2	3	5
4	1154	3	4	7
5	1213	9	4	13
6	1210	4	2	6
7	1199	9	3	12
8	1206	4	11	15
9	1194	1	11	12
10	1187	2	5	7
11	1184	0	5	5
12	1183	1	2	3
13	1179	4	0	4
14	1182	0	2	2
15	1182	6	5	11
TOTAL		**56**	**71**	**127**

Table 13. German artillery data. Dst+Abnd denotes destroyed and abandoned artillery. OH denotes the on hand amount.

D. COMPARISON OF MANPOWER AND WEAPON SYSTEMS

1. Personnel statistics

This section presents statistics on Soviet and German personnel strength and casualties during the campaign for the purpose of gaining insight about the Battle of Kursk.

a. On hand personnel

Figure 5 shows daily OH personnel for both forces in the southern front of the Battle of Kursk, as represented in the KDB. The OH total includes all personnel, attached to all combat units (line units and headquarter units through Army), both committed and uncommitted.

b. Personnel casualties

Figures 6 and 7 show daily and cumulative total casualties for both sides in the southern front of the Battle of Kursk. The total casualty includes all personnel,

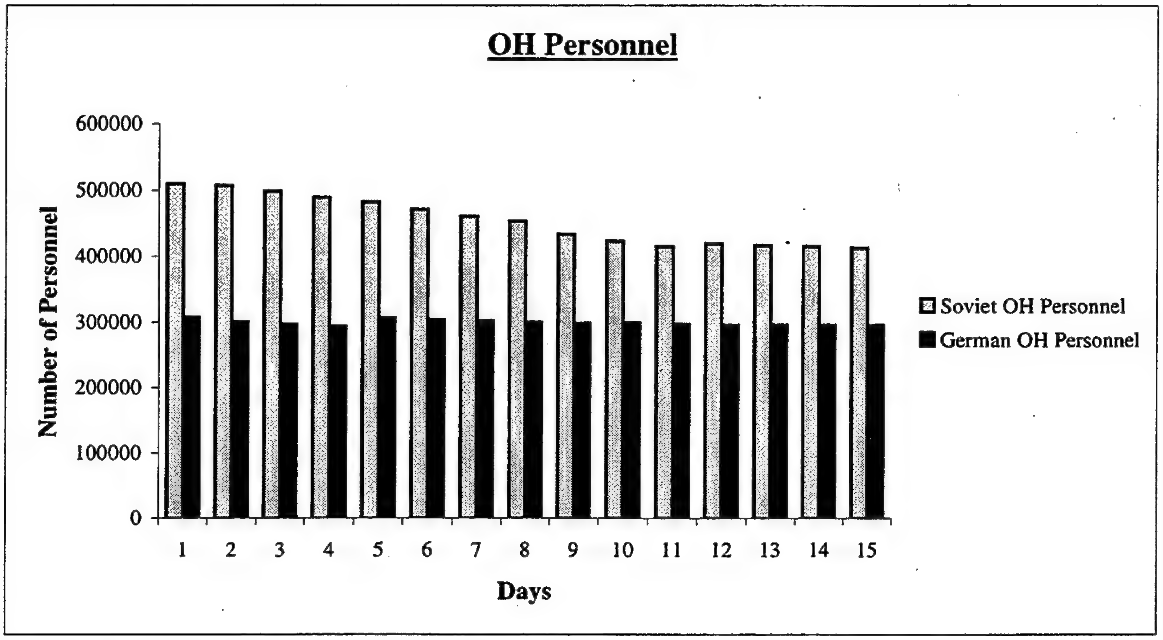

Figure 5. Daily total OH personnel for Soviet and German forces. Notice the steady decline in number of Soviet OH personnel until they counterattacked on July 12.

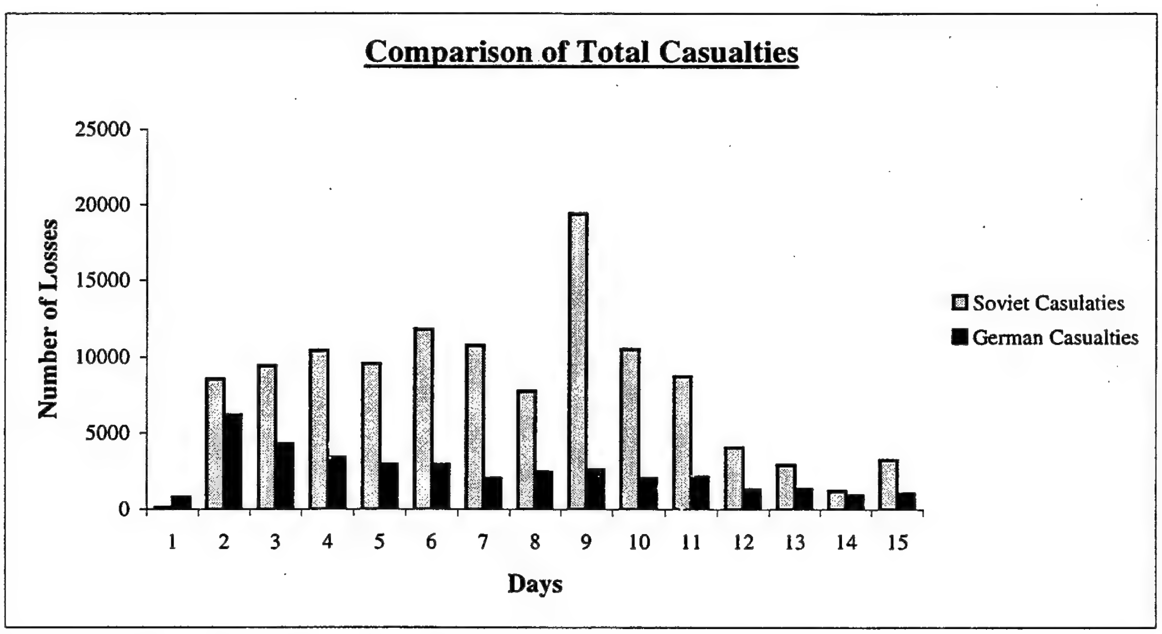

Figure 6. Daily total personnel casualty. Notice the great amount of casualties the Soviets had on July 12. Following this day, the battle lost its intensity for both sides.

both committed and uncommitted. When initial forces are considered, total casualties amounted to 23 percent of the initial Soviet force and 12 of the initial German force. Overall, the Soviets had three (3.24) casualties for every one German casualty.

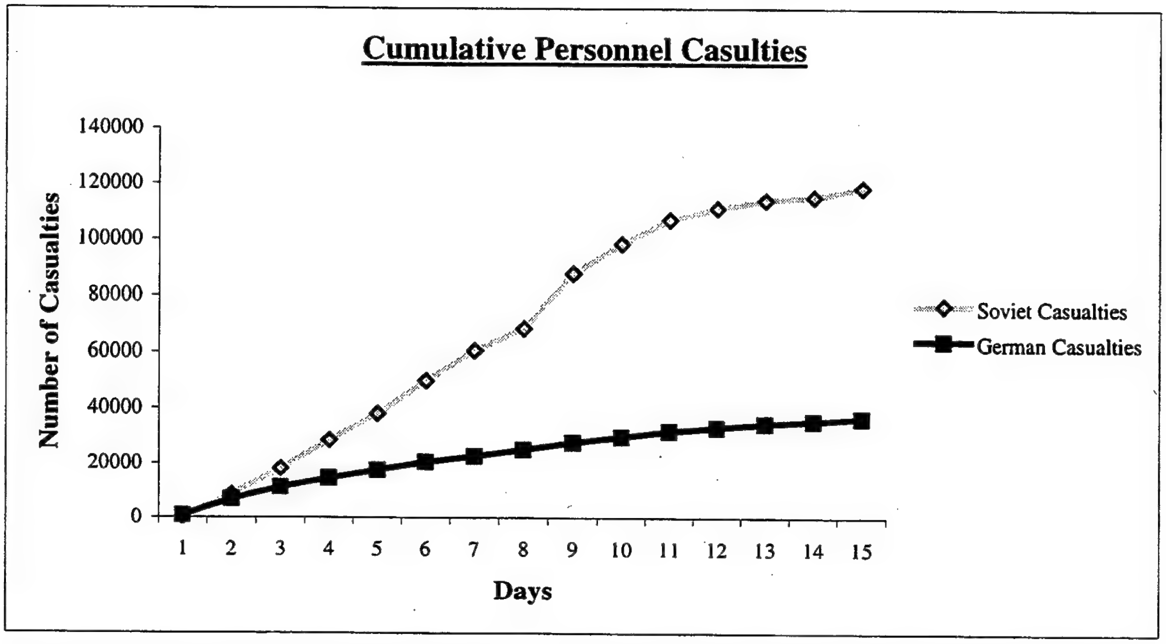

Figure 7. Comparison of daily cumulative number of total personnel casualties. There is a sharp increase in the number of Soviet casualties on July 12.

For more detailed information about the type of casualties, see Appendix A Part A.

2. Tank statistics

This section presents statistics on Soviet and German tank weapon system strength and losses during the campaign for the purpose of gaining insight about the Battle of Kursk.

a. On hand tanks

Figure 8 shows daily OH tanks for both sides in the southern front of the Battle of Kursk, as represented in the KDB. The number of OH tanks includes all tanks, both committed and uncommitted.

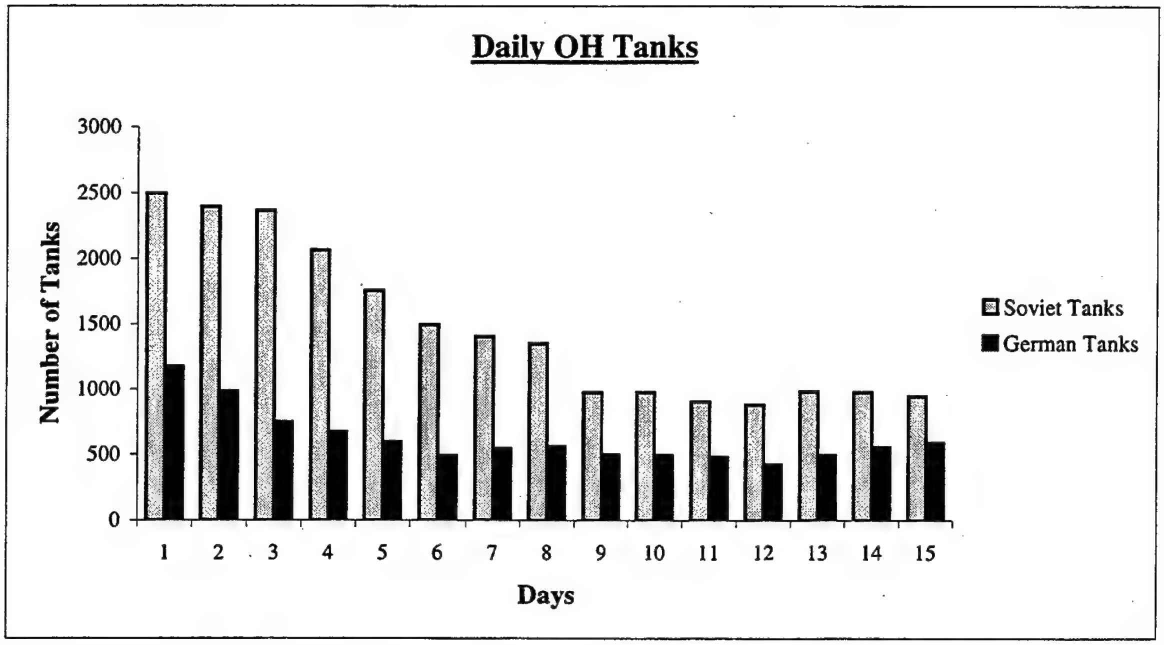

Figure 8. Comparison of daily number of total OH tanks. The Battle of Kursk was a major tank battle. Notice the sharp decline in the number of tanks for both sides in the first half of the battle.

b. Tank losses

Figures 9 and 10 show daily and cumulative total tank losses, respectively, for both sides in the southern front of the Battle of Kursk. When initial forces are considered, total tank losses amounted to 91 (0.910) percent of the initial amount of Soviet tanks and 104 (1.046) percent of the initial amount of German tanks (i.e. the Germans lost more tanks than their initial number of tanks). Overall, the Soviets lost almost 2 (1.84) tanks for every German tank lost.

3. Armored personnel carrier statistics

This section presents statistics on Soviet and German APC weapon system strength and losses during the campaign for the purpose of gaining insight about the Battle of Kursk.

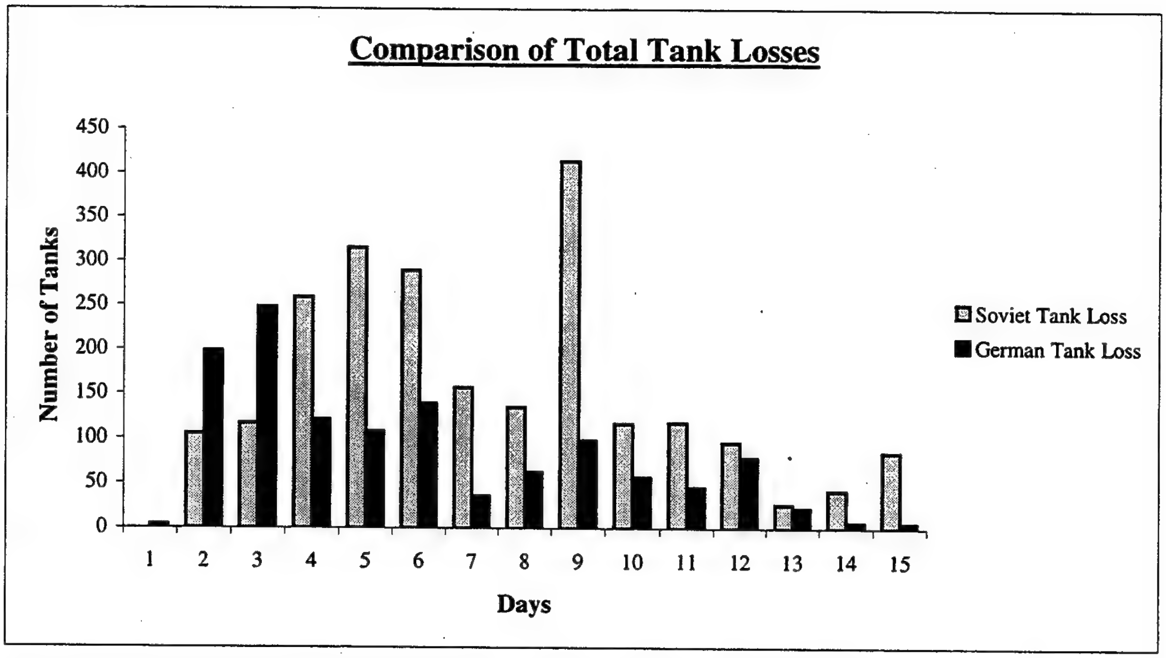

Figure 9. Comparison of daily number of tank losses. Notice the enormous number of Soviet tank loss on day 9.

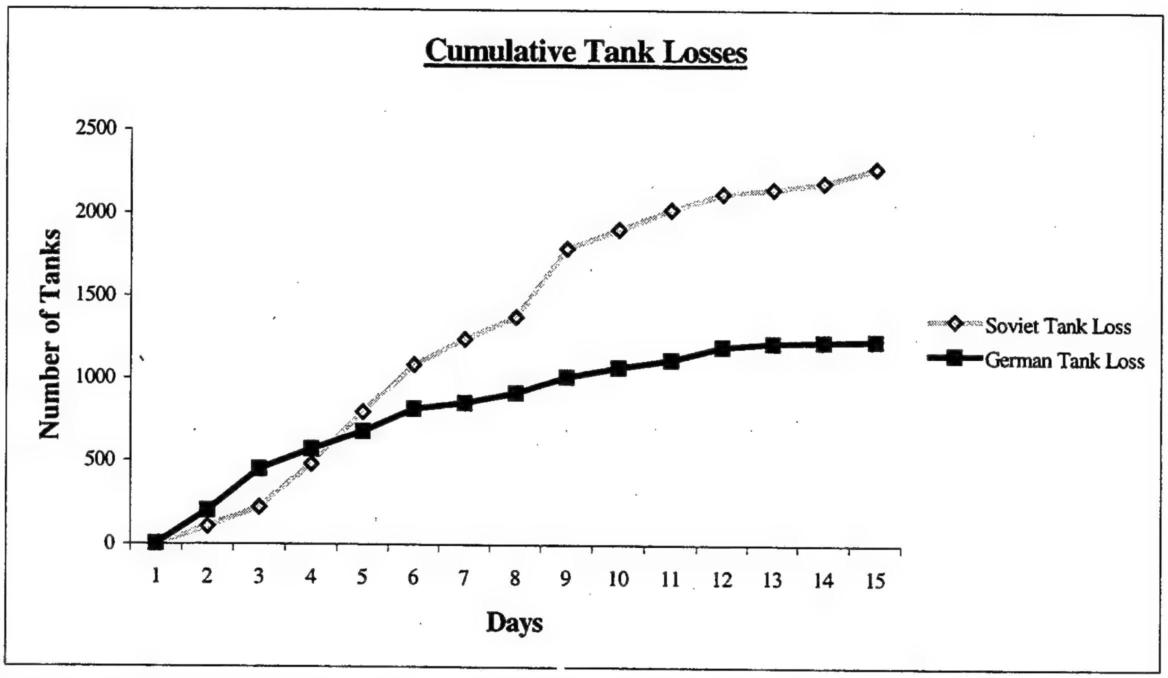

Figure 10. Comparison of daily cumulative number of tank losses. The Soviets lost almost 25% of their OH tanks on day 9.

a. On hand armored personnel carrier

Figure 11 shows daily OH APC for both sides in the southern front of the Battle of Kursk, as represented in the KDB. The number of OH APCs includes all APC, both committed and uncommitted.

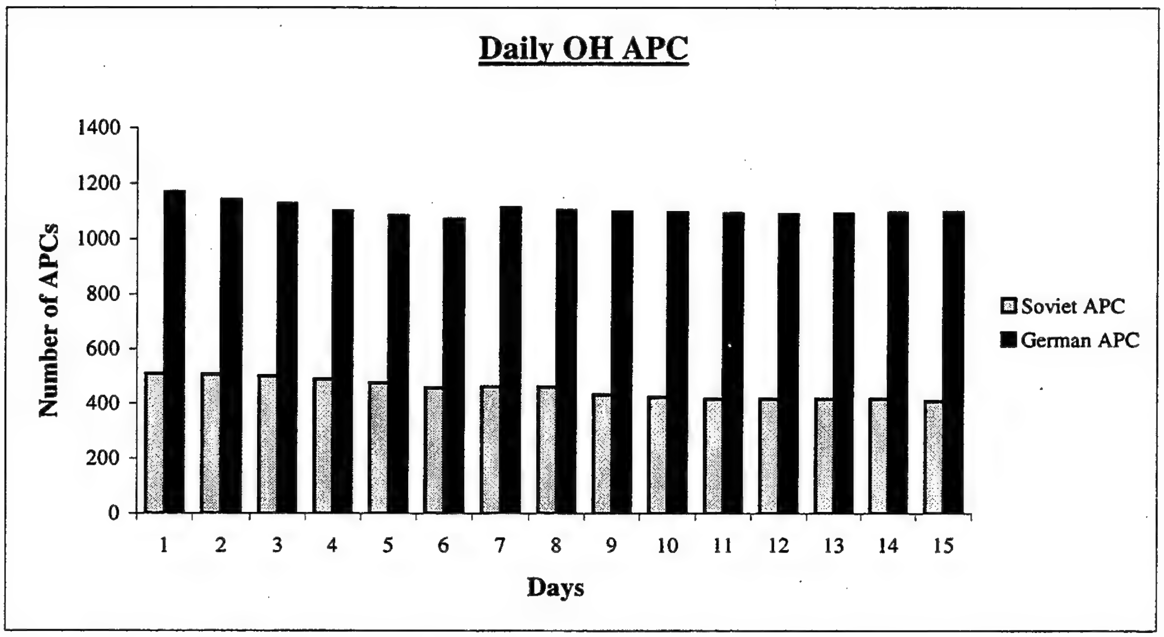

Figure 11. Comparison of daily number of total OH APCs. The number of Soviet OH APCs showed a general decline throughout the battle.

b. Armored personnel carrier losses

Figures 12 and 13 show daily and cumulative total APC losses, consecutively, for both sides in the southern front of the Battle of Kursk. When initial forces are considered, total APC losses amounted to 23 (22.89) percent of the initial amount of Soviet APCs and 16 (16.41) percent of the initial amount of German APCs. Overall, the Germans lost 1.64 APCs for every Soviet APC lost.

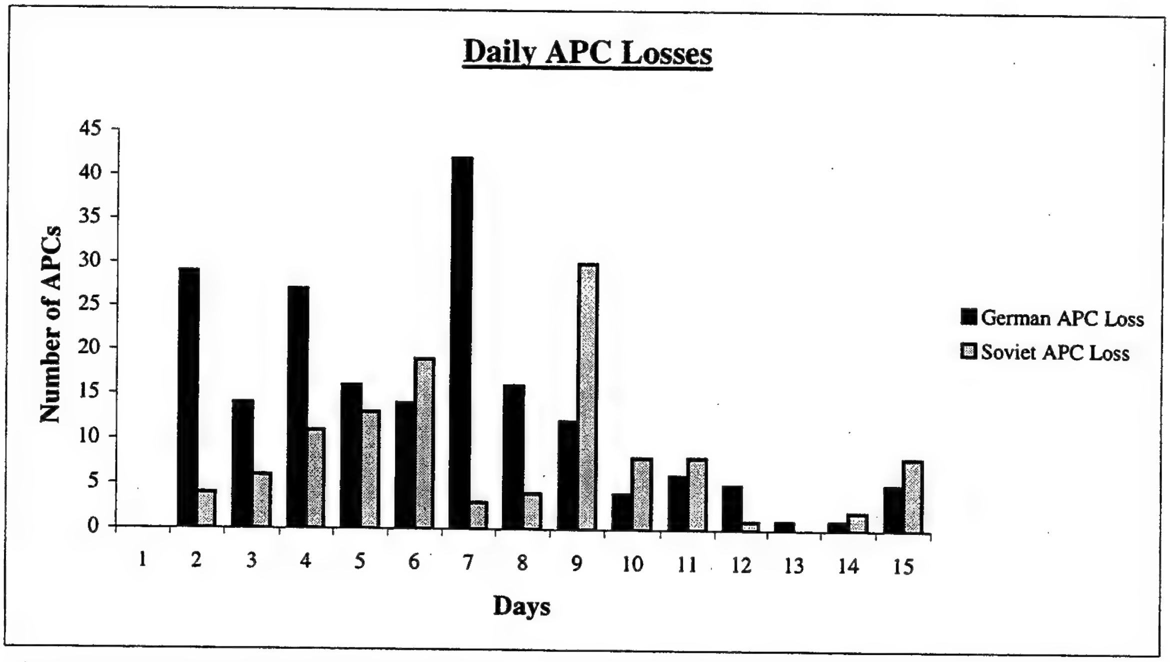

Figure 12. Comparison of daily number of APC losses. Notice the high number of German APC losses on day 7 and the high number of Soviet APC losses on day 9. Both sides did not lose any APCs on day 1.

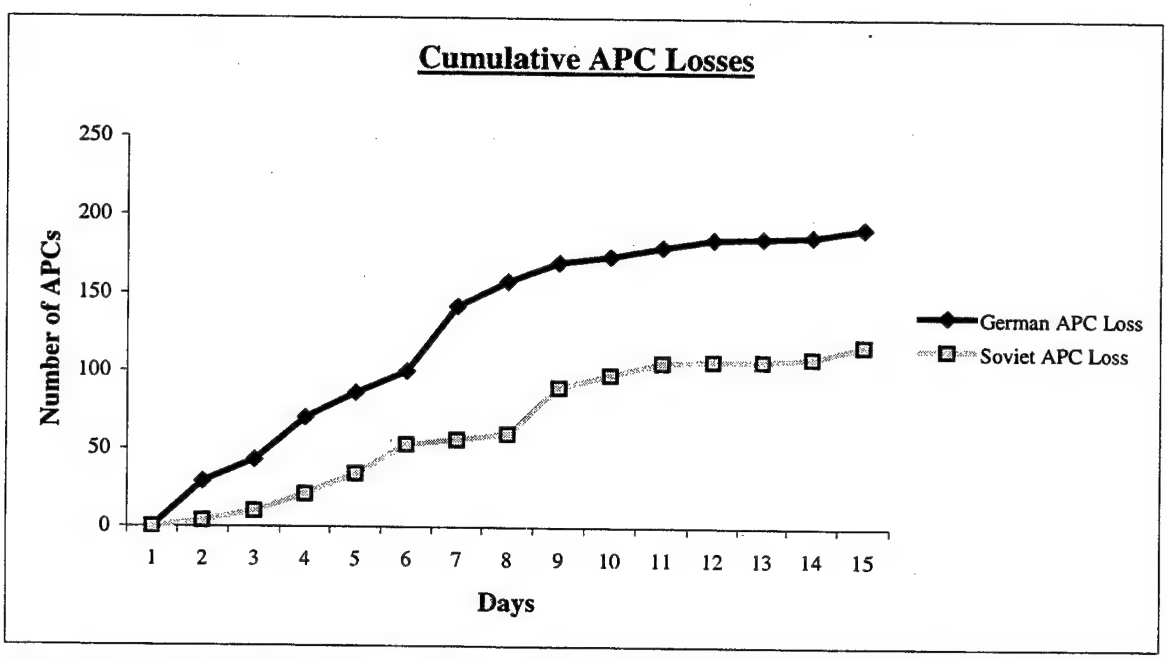

Figure 13. Comparison of daily cumulative number of APC losses. Throughout the battle, Germans lost more APCs than the Soviets did.

4. Artillery statistics

This section presents statistics on Soviet and German Artillery weapon system strength and losses during the campaign for the purpose of gaining insight about the Battle of Kursk.

a. On hand artillery

Figure 14 shows daily OH Artillery for both sides in the southern front of the Battle of Kursk, as represented in the KDB. The number of OH artillery includes all artillery, both committed and uncommitted.

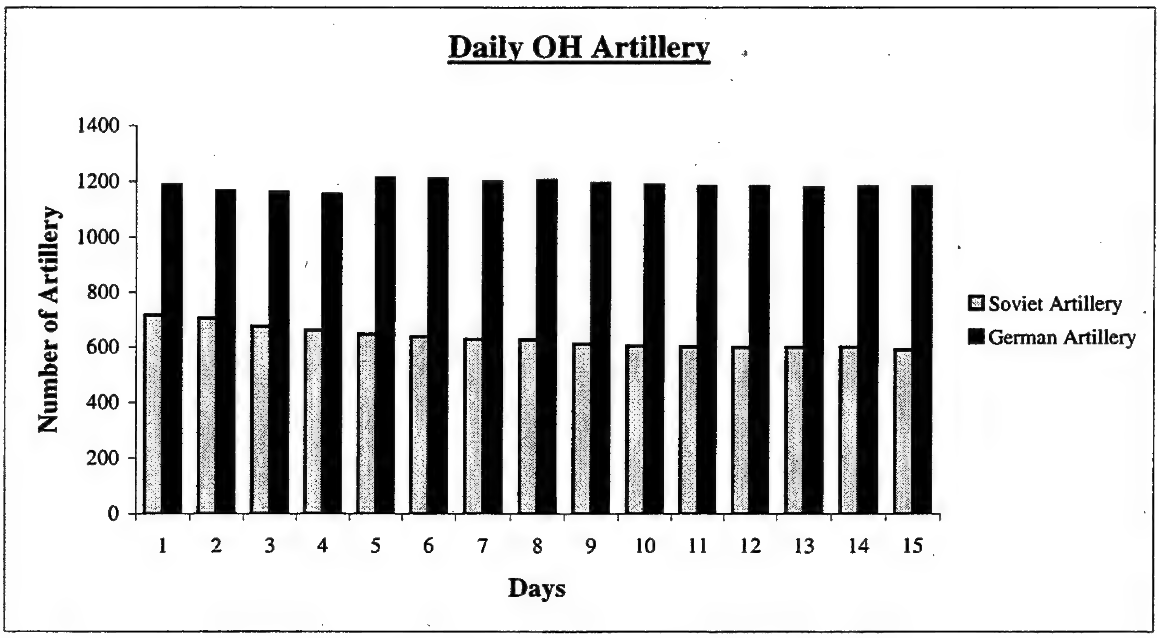

Figure 14. Comparison of daily number of total OH artillery. The number of German artillery was higher than the number of Soviet artillery throughout the battle.

b. Artillery losses

Figures 15 and 16 show daily and cumulative total artillery losses, consecutively, for both sides in the southern front of the Battle of Kursk. When initial forces are considered, total artillery losses amounted to 20 (0.200) percent of the initial

amount of Soviet artillery and 11 (0.106) percent of the initial amount of German artillery. Overall, the Soviets lost 1.13 artillery for every German artillery lost.

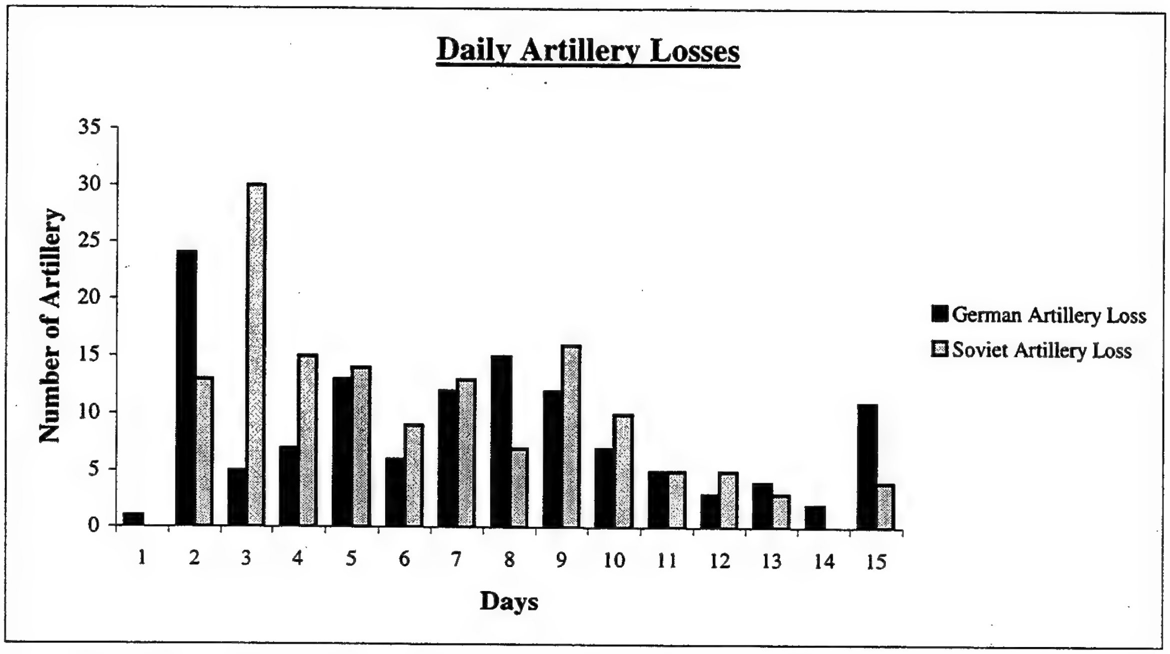

Figure 15. Comparison of daily number of artillery losses. German artillery loss was higher for the first few days of the battle. There were no Soviet losses on days 1 and 14.

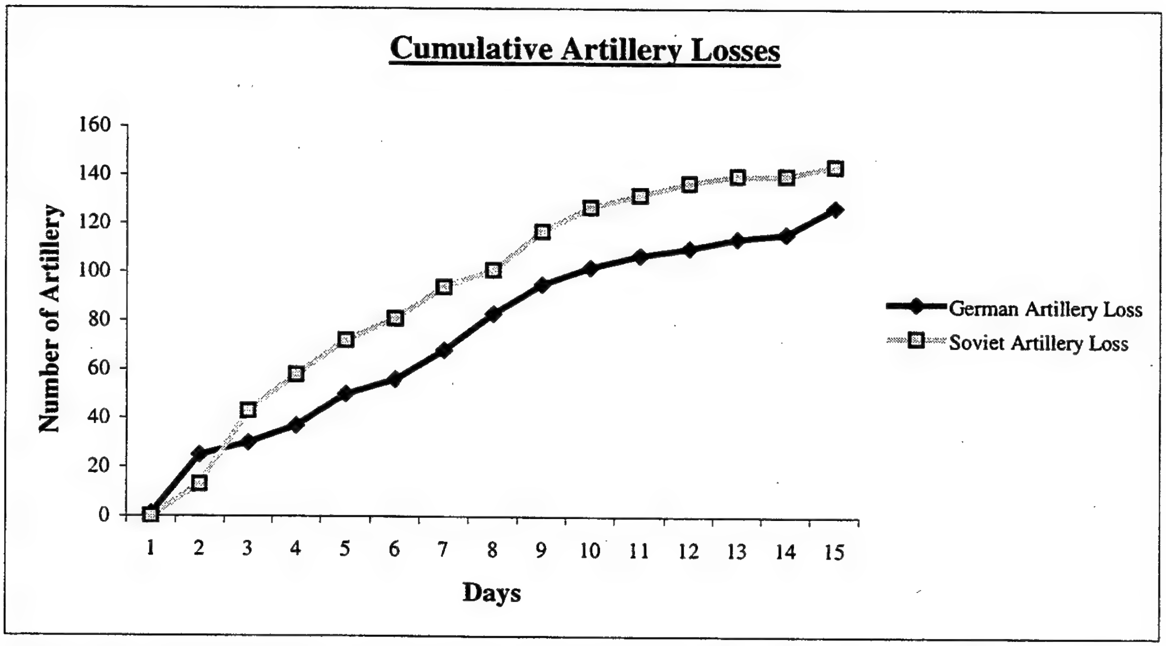

Figure 16. Comparison of daily cumulative number of artillery losses. Soviet artillery losses began to increase on the third day and remained higher throughout the battle.

For more detailed information about the type of losses for the tank, APC and artillery weapon systems, both for the Germans and Soviets, see Appendix A Part B, Part C and Part D, respectively.

5. Air sorties

The air sortie data given in the KOSAVE [Ref.12] report consists of the number of air-air role sorties, ground attack role sorties, reconnaissance role sorties and evacuation role sorties (used solely by Germans). The information on air sorties is given in Table 14.

Days	Soviet			German			
	air-air	grnd att.	recon.	air-air	grnd att.	recon.	evac.
1	143	1	15	64	160	0	0
2	1051	600	14	371	1942	74	67
3	778	613	20	253	1356	77	138
4	899	661	14	297	1499	91	117
5	707	669	62	229	1426	82	129
6	490	472	65	248	1286	99	83
7	322	383	71	116	530	44	81
8	410	348	25	176	809	66	70
9	501	603	67	191	460	57	107
10	406	623	58	204	451	55	136
11	593	704	46	238	1147	88	135
12	182	369	48	132	541	37	116
13	454	681	33	145	278	82	95
14	268	336	17	40	122	33	71
15	239	377	13	40	18	41	68
TOTAL	7443	7440	568	2744	12025	926	1413

Table 14. Number of air sorties for Soviets and Germans. Air-air denotes number of air-air role air sorties, grnd.att. denotes the number of ground attack role air sorties, recon. denotes the number of reconnaissance role air sorties and evac. denotes the number evacuation role air sorties. Evacuation role air sorties are used solely by Germans.

Figure 16 shows a comparison of the number of each type of air sorties for both sides for the Battle of Kursk.

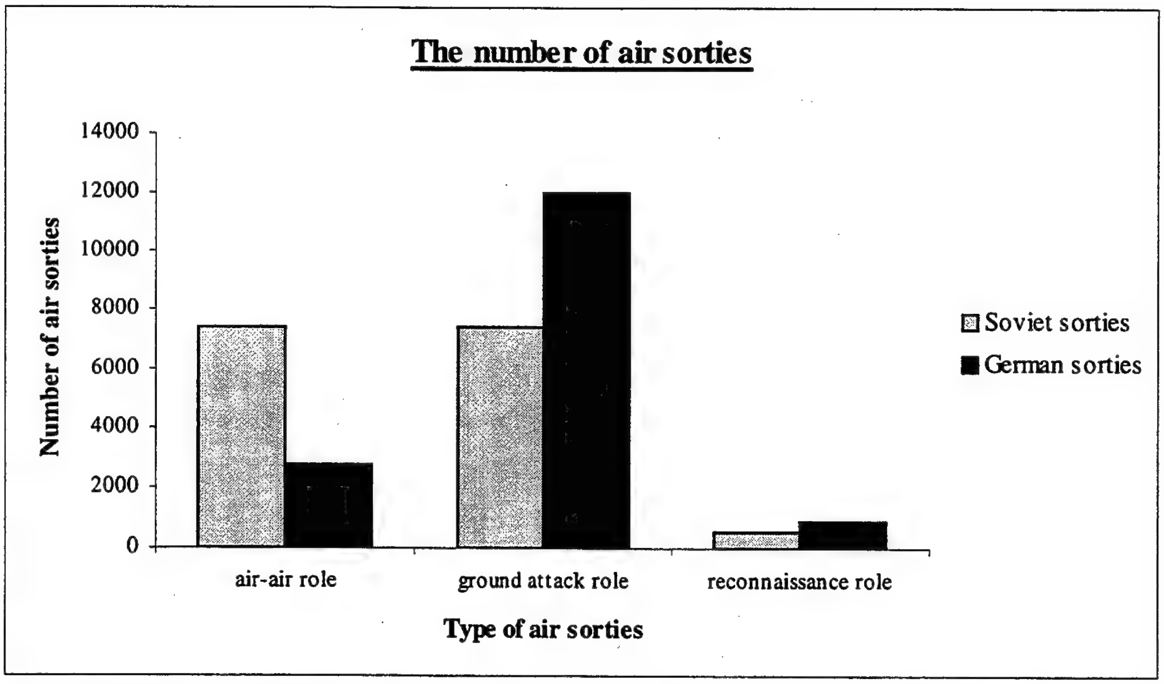

Figure 17. Number of each type of air sorties for Germans and Soviets. When the number of air-air role sorties are compared, the Soviets are superior to the Germans. When the number of ground attack role sorties are compared, Germans are superior to Soviets.

IV. COMPARATIVE AND EXPLORATIVE ANALYSIS OF BATTLE OF KURSK DATA

A. APPLICATION OF PREVIOUS STUDIES

1. Application of Bracken's methodology

This section analyzes the Battle of Kursk data following the same steps used by Bracken in his study, and subsequently applies Bracken's models to the Battle of Kursk data. The Battle of Kursk data is formatted and presented in tables using the same methodology, explained in detail in Chapter 3 and formatting techniques as Bracken did in his study. Tables 15 through 18 present data on combat manpower, APCs, tanks, and artillery consecutively for days 1-15 of the Battle of Kursk, from June 4, 1943 to June 18, 1943, for both the German and the Soviet forces.

Day	Blue manpower	Blue casualties	Red manpower	Red casualties
1	510252	130	307365	800
2	507698	8527	301341	6192
3	498884	9423	297205	4302
4	489175	10431	293960	3414
5	481947	9547	306659	2942
6	470762	11836	303879	2953
7	460808	10770	302014	2040
8	453126	7754	300050	2475
9	433813	19422	298710	2612
10	423351	10522	299369	2051
11	415254	8723	297395	2140
12	419374	4076	296237	1322
13	416666	2940	296426	1350
14	415461	1217	296350	949
15	413298	3260	295750	1054

Table 15. Combat manpower data for both sides. Casualties are killed, wounded, captured/missing in action, and disease and nonbattle injuries. Notice the low casualty rates for day 1, when the offensive had not really started.

Day	Blue APCs	Blue APCs Killed	Red APCs	Red APCs killed
1	511	0	1170	0
2	507	4	1142	29
3	501	6	1128	14
4	490	11	1101	27
5	477	13	1085	16
6	458	19	1073	14
7	463	3	1114	42
8	462	4	1104	16
9	432	30	1099	12
10	424	8	1096	4
11	418	8	1093	6
12	417	1	1089	5
13	417	0	1092	1
14	417	2	1095	1
15	409	8	1098	5

Table 16. APC data for both sides. Killed are destroyed+abandoned and damaged. Notice the high number of German APC losses on day 7 and the high number of Soviet APC losses on day 9.

Day	Blue Tanks	Blue Tanks Killed	Red Tanks	Red Tanks killed
1	2500	0	1178	4
2	2396	105	986	198
3	2367	117	749	248
4	2064	259	673	121
5	1754	315	596	108
6	1495	289	490	139
7	1406	157	548	36
8	1351	135	563	63
9	977	414	500	98
10	978	117	495	57
11	907	118	480	46
12	883	96	426	79
13	985	27	495	23
14	978	42	557	7
15	948	85	588	6

Table 17. Tank data for both sides. Killed are destroyed+abandoned and damaged. Notice the decrease in the number of tank losses after day 9. After day 9, the battle lost its intensity.

Table 19 presents data on total forces, where the data from Tables 16-18 on combat manpower, APCs, tanks, and artillery are weighted by 1, 5, 40, and 20, respectively. Bracken [Ref.8] states in his study that, "The weights given above are

Day	Blue artillery	Blue artillery killed	Red artillery	Red artillery killed
1	718	0	1189	1
2	705	13	1166	24
3	676	30	1161	5
4	661	15	1154	7
5	648	14	1213	13
6	640	9	1210	6
7	629	13	1199	12
8	628	7	1206	15
9	613	16	1194	12
10	606	10	1187	7
11	603	5	1184	5
12	601	5	1183	3
13	600	3	1179	4
14	602	0	1182	2
15	591	4	1182	11

Table 18. Artillery data for both sides. Killed are destroyed+abandoned and damaged. On the initial days of the battle, German artillery losses were higher than the Soviet artillery losses.

consistent with those of studies and models of the U.S. Army Concepts Analysis Agency. Virtually all theater-level dynamic combat simulation models incorporate similar weights, either as inputs or as decision parameters computed as the simulations progress".

Day	Blue forces	Blue losses	Red forces	Red losses
1	591527	130	384335	920
2	586353	11167	373411	11257
3	575769	12993	364265	9532
4	559345	16266	359085	6249
5	545332	16472	372524	5702
6	528552	18071	367444	6043
7	516403	14445	366504	3450
8	507576	10754	365070	4415
9	480033	28492	361965	5112
10	469271	13302	362229	3491
11	459604	11323	359820	3290
12	463159	6201	357522	3047
13	462451	3600	358946	1975
14	461186	2067	360245	1174
15	457943	5160	360280	1639

Table 19. Data on aggregated forces. Forces are combat manpower, APCs, tanks and artillery which are weighted by 1, 5, 20 and 40 respectively. The number of Soviet losses on day 9 is almost three times the amount of Soviet loss on day 8.

a. Estimation of Parameters

The parameters of the model are chosen to minimize the sum of squared residuals between the estimated and actual attrition. Using 15 days of the Battle of Kursk data, where the first 8 days the Germans attack and the last 7 days the Soviets attack, it is desired to minimize:

$$\text{SSR} = \sum_{n=1}^{8}(\dot{B}_n - adR_n^p B_n^q)^2 + \sum_{n=1}^{8}(\dot{R}_n - b(1/d)B_n^p R_n^q)^2$$

$$+ \sum_{n=9}^{15}(\dot{B}_n - a(1/d)R_n^p B_n^q)^2 + \sum_{n=9}^{15}(\dot{R}_n - bdB_n^p R_n^q)^2 \qquad (5)$$

where n denotes the index for the 15 days of the battle. Using the data given in Table 17, the above procedure will give a different SSR value for each set of parameters, i.e. combination of a, b, p, q and d values. It will evaluate $\text{SS}(a_i, b_j, p_k, q_l, d_m)$ for all combinations of i, j, k, l and m where i=1,....,9, j=1,....,9, k=1,....,21, l=1,....,21, and m=1,....,9.

The range of possibilities allowed for the parameters, for the model with the tactical parameter d will be:

$$(a_1, \ldots\ldots\ldots, a_9) = (4\times10^{-9}, \ldots\ldots\ldots\ldots, 1.2\times10^{-8}),$$

$$(b_1, \ldots\ldots\ldots, b_9) = (4\times10^{-9}, \ldots\ldots\ldots\ldots, 1.2\times10^{-8}),$$

$$(p_1, \ldots\ldots\ldots, p_{21}) = (0.0, \ldots\ldots\ldots, 2.0),$$

$$(q_1, \ldots\ldots\ldots, q_{21}) = (0.0, \ldots\ldots\ldots, 2.0),$$

$$(d_1, \ldots\ldots\ldots, d_9) = (0.6, \ldots\ldots\ldots, 1.4).$$

These ranges were selected by Bracken himself.

There are a total of 9x9x21x21x9 = 321489 combinations of the estimated parameters. The algorithm searches all combinations and determines the parameters that minimize the sum of squared residuals for the data given in Table 17 as:

$$SS(a_9, b_6, p_2, q_{21}, d_4) = 8.65 \times 10^8$$

with the estimated parameters of:

$$a_9 = 1.2 \times 10^{-8}, b_6 = 9 \times 10^{-9}, p_2 = 0.1, q_{21} = 2.0, d_4 = 0.9.$$

Notice that the values of the *a* parameter and the *q* parameter are on the boundary.

Now, considering the estimation of parameters for the model without the tactical parameter *d*, the ranges of possibilities allowed will be the same as those in the previous procedure, except for the tactical parameter *d*. There are now a total of 9x9x21x21 = 35721 combinations of parameters. The algorithm searches all combinations and determines the parameters that minimize the sum of squared residuals for the data given in Table 17 as:

$$SS(a_9, b_6, p_4, q_{19}) = 8.88 \times 10^8$$

with the estimated parameters of:

$$a_9 = 1.2 \times 10^{-8}, b_6 = 9 \times 10^{-9}, p_2 = 0.3, q_{21} = 1.8.$$

Table 20 gives the sums of squared residuals for different values of *d,* and shows which *a, b, p, q* combinations gives the minimum sums of squared residuals for the various *d* values. Table 20 also shows the sensitivity of the *p* and *q* parameters to the *d* parameter and suggests that the sums of the squared residuals are similar within a wide range of parameter values.

b. *Results*

The best fitting results for the two models for the Battle of Kursk data are:

Bracken's model 1 with tactical parameter *d*,

$$\dot{B} = 1.2\times10^{-8}(\frac{10}{9} \text{ or } \frac{9}{10})R^{0.1}B^{2.0} \tag{6}$$

$$\dot{R} = 9\times10^{-9}(\frac{9}{10} \text{ or } \frac{10}{9})B^{0.1}R^{2.0} \tag{7}$$

d	SSR	*a*	*b*	*p*	*q*
0.5	1.38E+9	9.00E-9	6.00E-9	0.1	2.0
0.6	1.15E+9	1.00E-8	7.00E-9	0.1	2.0
0.7	9.84E+8	1.20E-8	8.00E-9	0.1	2.0
0.8	8.87E+8	1.20E-8	9.00E-9	0.1	2.0
0.9	8.65E+8	1.20E-8	9.00E-9	0.1	2.0
1.0	8.88E+8	1.20E-8	9.00E-9	0.3	1.8
1.1	9.34E+8	1.20E-8	8.00E-9	0.5	1.6
1.2	9.90E+8	1.20E-8	7.00E-9	0.7	1.4
1.3	1.05E+9	1.20E-8	7.00E-9	0.8	1.3
1.4	1.10E+9	1.20E-8	6.00E-9	1.0	1.1
1.5	1.16E+9	1.20E-8	5.00E-9	1.2	0.9
1.6	1.21E+9	1.20E-8	5.00E-9	1.3	0.8
1.7	1.25E+9	1.20E-8	4.00E-9	1.5	0.6

Table 20. SSR values for different *d* values. *a* and *b* values are varied between 8×10^{-9} and 1.2×10^{-8} with increments of 1×10^{-9}, *p* and *q* values are varied between 0.0 and 2.0 with increments of 0.1. The lowest SSR value is observed to be 8.65E+8 when *d*=0.9.

Bracken's model 3 without the tactical parameter *d*

$$\dot{B} = 1.2\times10^{-8}R^{0.3}B^{1.8} \tag{8}$$

$$\dot{R} = 9\times10^{-9}B^{0.3}R^{1.8} \tag{9}$$

The parameters found in equations IV.A.1.b.(6), IV.A.1.b.(7), IV.A.1.b.(8), IV.A.1.b.(9) suggest that one side's losses are more a function of his own forces rather than a function of the opponent's forces. This result is similar to what Fricker found in his study. There are boundaries set for the search of parameters that

give the best fit. There may be other sets of parameters that are out of the range of possibilities allowed by this method, and they may give a better fit for the data. The fact that some of the best fitting parameters are on the boundary supports this hypothesis.

Figures 18 and 19 show the real and fitted values found using the model with the d parameter (i.e., using equations IV.A.1.b.(6) and IV.A.1.b.(7), for the Soviet and the German forces respectively). Figures 20 and 21 show the real and fitted values found using the model without the d parameter (i.e. using the formulas IV.A.1.b.(8) and IV.A.1.b.(9), for the Soviet and German forces, respectively).

When the plots given in Figures 18 and 20 are examined, there appears to be three phases in the battle. It is also apparent that the battle lost its intensity after July 12. The model underestimates the casualties for the beginning part and the last part of the battle while overestimating the 8 days in a row between these two periods. This pattern suggests that fitting a model with change points may improve the fit to the data.

For the model with the tactical parameter, p-q=-1.9, and for the model without the tactical parameter p-q=-1.5. These two results imply that the Battle of Kursk data does not fit any one of the basic Lanchester linear, square or logarithmic laws.

For both cases, parameters a and b are significantly small and $a > b$. This suggests that individual German effectiveness was greater than individual Russian effectiveness.

For the purpose of comparing a variety of models throughout this thesis, R^2 values are also computed together with the SSR for each model, where R^2 is given as:

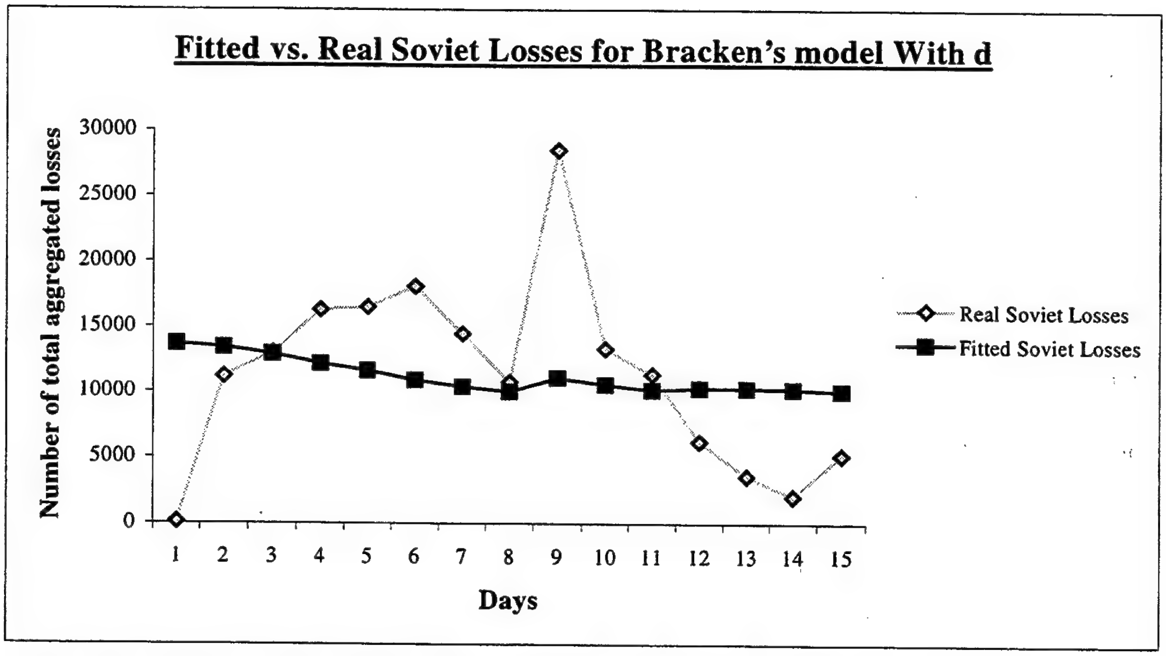

Figure 18. Fitted Soviet losses found by using Bracken's model with parameter *d*, plotted versus real Soviet losses. Notice the three-phase pattern in the model's fit to the battle data where the model overestimates the first two days and the last four days of the battle while underestimating the part between these two phases.

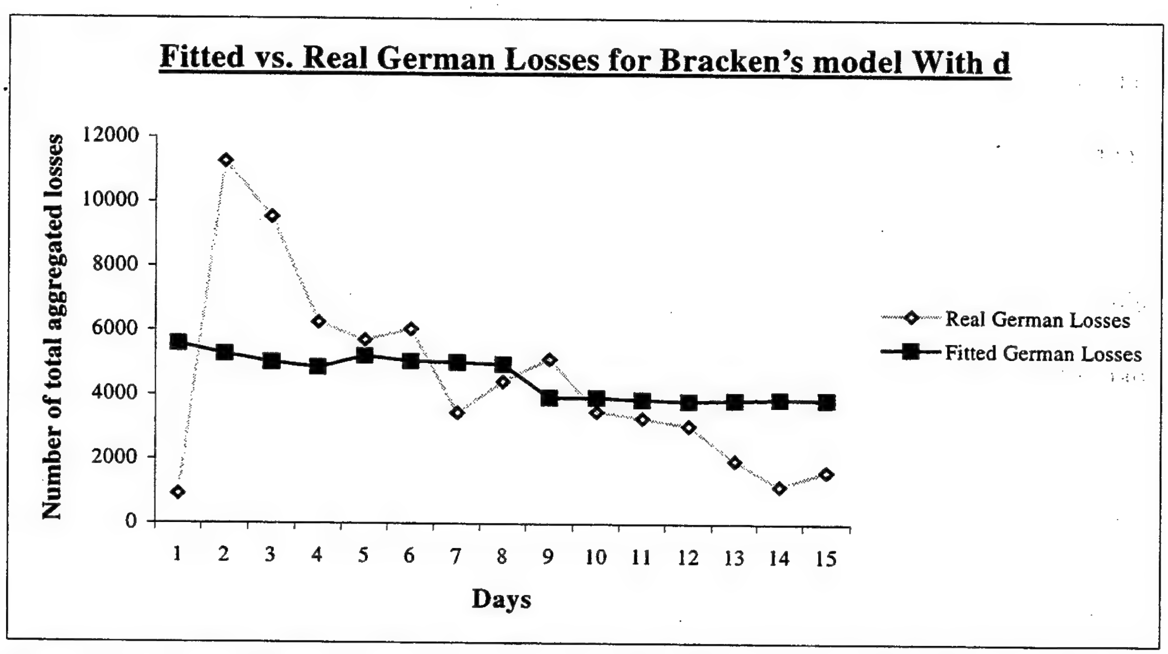

Figure 19. Fitted German losses found by using Bracken's model with parameter *d*, plotted versus real German losses. After the Soviets went into offense, the battle was not as intense.

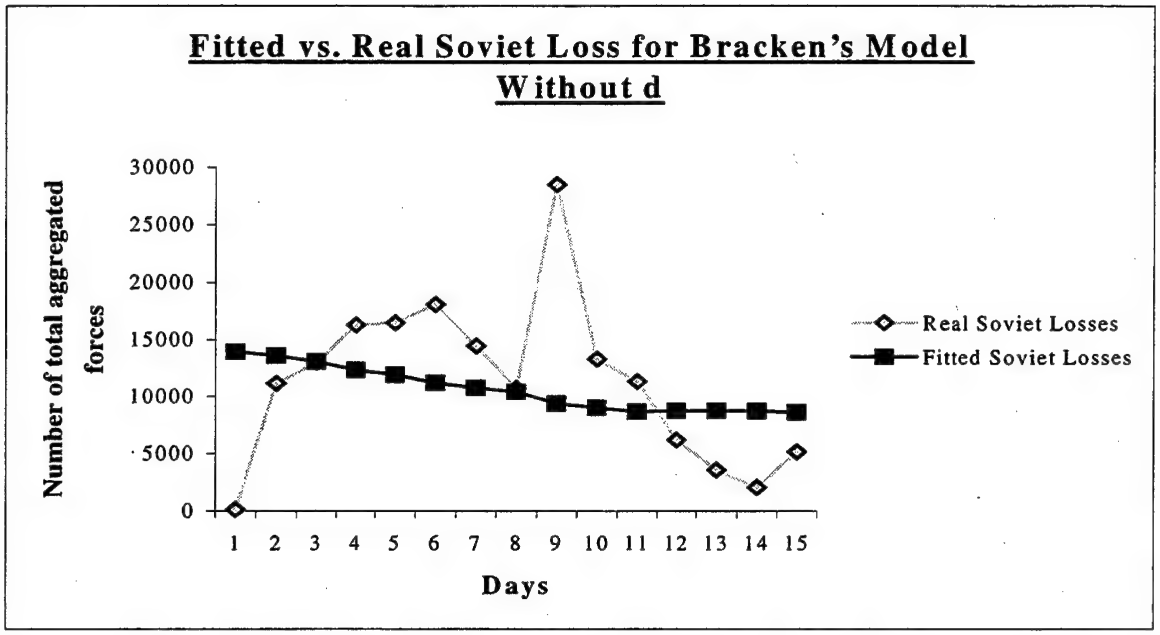

Figure 20. Fitted Soviet losses found by using Bracken's model without parameter d, plotted versus real Soviet losses. Like the plot given in Figure 18, notice the three-phase pattern in the model's fit to the battle where the model overestimates the first two days and the last four days of the battle while underestimating the part between these two phases.

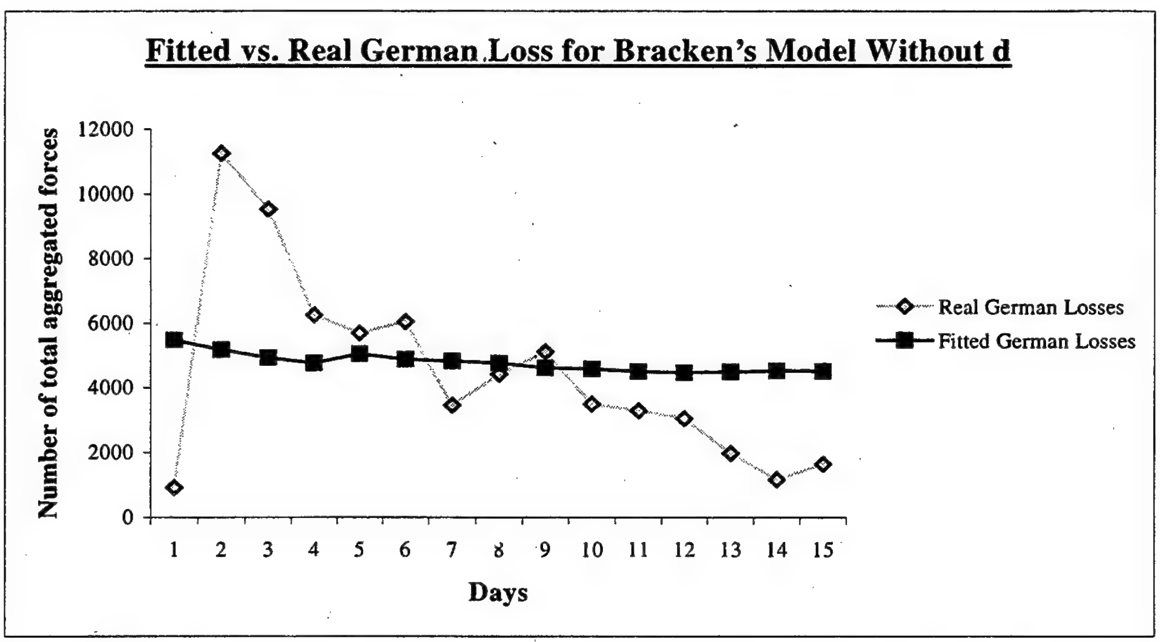

Figure 21. Fitted German losses found by using Bracken's model without parameter *d*, plotted versus real German losses. After the Soviets went into offense, the battle was not as intense.

$$R^2 = 1 - \frac{SSR}{SST} = 1 - \frac{\sum_i (Y - \hat{Y})^2}{\sum_i (Y - \overline{Y})^2} \tag{10}$$

where $\hat{Y}$, Y and $\overline{Y}$ denote the estimated value, the real value and the mean value of the Y parameter (daily casualties) which are indexed by days. A greater R^2 value indicates a better fit. It is possible to get a negative R^2 value, implying that the fitted model yields worse results than using the average daily losses as an estimate.

Table 21 shows the results for Bracken's models as a whole.

Name of the model	***a***	***b***	***p***	***q***	***d***	***SSR***	R^2
Bracken Model 1 Ardennes	8.0E-9	1.0E-8	1.0	1.0	1.25	1.63E+7	0.2552
Bracken Model 3 Ardennes	8.0E-9	1.0E-8	1.3	0.7	1.0	2.08E+8	0.0493
Bracken Model 1 Kursk	1.2E-8	9.0E-9	0.1	2.0	0.9	8.65E+8	0.0006
Bracken Model 3 Kursk	1.2E-8	9.0E-9	0.3	1.8	1.0	8.88E+8	-0.0266

Table 21. Bracken's results for his models with the tactical parameter d for both the Ardennes and Kursk data.

Upon examination of the fits of Bracken's models found in this section, it is clear that the battle did not start until the second day. Thus, the first day of data was dropped in fitting the data to the models in the rest of the thesis. More detailed explanation on this approach is given in Section IV.B.1.

Bracken's Model 1 was refit using only the last 14 days of the data. The new parameter estimates are: a=1.2×10^{-8}, b=1.0×10^{-8}, p=0.1, q=2.0, d=1.0. The SSR value dropped to 6.50×10^{8} and the R^2 value increased to 0.0919.

2. Application of Fricker's methodology

In his study, Fricker presents an alternate way to structure the data that reflects the effects of both previous casualties and incremental reinforcements. His idea is based on the fact that the casualties occur according to the Lanchester equations that use the previous day's force size, and for any given day, the previous day's force size also depends on the transfer of troops in or out of the fighting force.

Because of this phenomenon, Fricker uses the following algorithm in his study to estimate the original total for each resource. The algorithm works sequentially stepping through the whole battle from day 1 to the last day of the battle. By using this algorithm, local reserves (X_{lr}) or the addition of reinforcements (X_r) are accounted for. The algorithm first uses local reserves for any force increase before using reinforcements. As described in Fricker's study [Ref.6], the algorithm is:

For resource *X*:

1. Set $X_r(t) = X_{lr}(t) = 0$

2. Let t=1:

   - If $X(t+1) > X(t) - \dot{X}(t)$ and $X_{lr} = 0$, then

     $X_r(t) = X_r(t) + [X(t+1) - X(t) + \dot{X}(t)]$

   - Else, if $X(t+1) > X(t) - \dot{X}(t)$ and $X_{lr} \geq X(t+1) - X(t) + \dot{X}(t)$,

     then $X_{lr}(t) = X_{lr}(t) - [X(t+1) - X(t) + \dot{X}(t)]$

   - Else, if $X(t+1) > X(t) - \dot{X}(t)$ and

     $0 < X_{lr}(t) < X(t+1) - X(t) + \dot{X}(t)$, then

     $X_r(t) = X_r(t) + [X(t+1) - X(t) + \dot{X}(t)] - X_{lr}(t), X_{lr}(t) = 0$.

- Else, if $X(t+1) < X(t) - \dot{X}(t)$, then

$$X_{lr}(t) = X_{lr}(t) + [X(t) - \dot{X}(t) - X(t+1)]$$

3. If $t < 31$, increment t and go to step #2; else $\tilde{X}(0) = X(0) + X_r(t)$

Then the new daily resources $\tilde{X}(t+1)$ are calculated as:

$$\tilde{X}(t+1) = \tilde{X}(t) - \dot{X}(t) \qquad t = 0,\ldots\ldots\ldots,31 \tag{11}$$

After the data is reformatted using the algorithm given above, Fricker applies linear regression to logarithmically transformed Lanchester equations for estimating the model parameters. After the logarithmic transformation, the basic Lanchester equations, given in I.B.(1) and I.B.(2), will look like:

$$\ln(\dot{B}) = \ln(a) + p\ln(R) + q\ln(B) \tag{12}$$

$$\ln(\dot{R}) = \ln(b) + p\ln(B) + q\ln(R) \tag{13}$$

Below is the Battle of Kursk data reformatted using Fricker's approach. For reformatting the data, the algorithm, which is explained in detail above, is applied to the given Battle of Kursk data.

Tables 22 and 23 present the raw manpower and weapon systems data, respectively. Tables 24 and 25 present the resulting reformatted Kursk data for manpower and weapon systems, respectively. Table 26 presents the aggregated force (except the first day) found by aggregating the data given in Tables 24 and 25.

The air sortie data given in the KOSAVE study [Ref.12] consists of the number of air-air role sorties, ground attack role sorties, reconnaissance role sorties and evacuation role sorties (solely used by Germans). Table 27 presents data on number of ground attack role sorties. Table 28 presents the aggregated force, after the air sortie data is added

(except the first day) by using the weight coefficient of 30, as used by Fricker, (i.e., the number of air sorties presented in Table 27 is multiplied by 30 and added to the aggregated force levels given in Table 26 to get the data presented in Table 28).

Day	Blue Available	Blue Killed	Red Available	Red Killed
1	510252	130	307365	800
2	507698	8527	301341	6192
3	498884	9423	297205	4302
4	489175	10431	293960	3414
5	481947	9547	306659	2942
6	470762	11836	303879	2953
7	460808	10770	302014	2040
8	453126	7754	300050	2475
9	433813	19422	298710	2612
10	423351	10522	299369	2051
11	415254	8723	297395	2140
12	419374	4076	296237	1322
13	416666	2940	296426	1350
14	415461	1217	296350	949
15	413298	3260	295750	1054

Table 22. Battle of Kursk manpower data for the Soviet and German forces.

Day	BLUE						RED					
	Available			Killed			Available			Killed		
	Tank	APC	Art.	Tank	APC	Art.	Tank	APC	Art.	Tank	APC	Art.
1	2500	511	718	0	0	0	1178	1170	1189	4	0	1
2	2396	507	705	105	4	13	986	1142	1166	198	29	24
3	2367	501	676	117	6	30	749	1128	1161	248	14	5
4	2064	490	661	259	11	15	673	1101	1154	121	27	7
5	1754	477	648	315	13	14	596	1085	1213	108	16	13
6	1495	458	640	289	19	9	490	1073	1210	139	14	6
7	1406	463	629	157	3	13	548	1114	1199	36	42	12
8	1351	462	628	135	4	7	563	1104	1206	63	16	15
9	977	432	613	414	30	16	500	1099	1194	98	12	12
10	978	424	606	117	8	10	495	1096	1187	57	4	7
11	907	418	603	118	8	5	480	1093	1184	46	6	5
12	883	417	601	96	1	5	426	1089	1183	79	5	3
13	985	417	600	27	0	3	495	1092	1179	23	1	4
14	978	417	602	42	2	0	557	1095	1182	7	1	2
15	948	409	591	85	8	4	588	1098	1182	6	5	11

Table 23. Battle of Kursk data for tanks, APCs, and artillery of the Soviet and German forces.

Day	Blue Available	Blue Killed	Red Available	Red Killed
1	529562	130	331292	800
2	529432	8527	330492	6192
3	520905	9423	324300	4302
4	511482	10431	319998	3414
5	501051	9547	316584	2942
6	491504	11836	313642	2953
7	479668	10770	310689	2040
8	468898	7754	308649	2475
9	461144	19422	306174	2612
10	441722	10522	303562	2051
11	431200	8723	301511	2140
12	422477	4076	299371	1322
13	418401	2940	298049	1350
14	415461	1217	296699	949
15	414244	3260	295750	1054

Table 24. The reformatted Battle of Kursk manpower data for the Soviet and German forces.

Day	BLUE						RED					
	Available			Killed			Available			Killed		
	Tank	APC	Art.	Tank	APC	Art.	Tank	APC	Art.	Tank	APC	Art.
1	3139	524	742	0	0	0	1815	1285	1298	4	0	1
2	3139	524	742	105	4	13	1811	1285	1297	198	29	24
3	3034	520	729	117	6	30	1613	1256	1273	248	14	5
4	2917	514	699	259	11	15	1365	1242	1268	121	27	7
5	2658	503	684	315	13	14	1244	1215	1261	108	16	13
6	2343	490	670	289	19	9	1136	1199	1248	139	14	6
7	2054	471	661	157	3	13	997	1185	1242	36	42	12
8	1897	468	648	135	4	7	961	1143	1230	63	16	15
9	1762	464	641	414	30	16	898	1127	1215	98	12	12
10	1348	434	625	117	8	10	800	1115	1203	57	4	7
11	1231	426	615	118	8	5	743	1111	1196	46	6	5
12	1113	418	610	96	1	5	697	1105	1191	79	5	3
13	1017	417	605	27	0	3	618	1100	1188	23	1	4
14	990	417	602	42	2	0	595	1099	1184	7	1	2
15	948	415	602	85	8	4	588	1098	1182	6	5	11

Table 25. The reformatted Battle of Kursk equipment data for tanks, APCs, and artillery of the Soviet and German forces.

Day	Blue forces	Blue losses	Red forces	Red losses
1	624512	11167	425017	11257
2	613345	12993	413760	9532
3	600352	16266	404228	6249
4	584086	16472	397979	5702
5	567614	18071	392277	6043
6	549543	14445	386234	3450
7	535098	10754	382784	4415
8	524344	28492	378369	5112
9	495852	13302	373257	3491
10	482550	11323	369766	3290
11	471227	6201	366476	3047
12	465026	3600	363429	1975
13	461426	2067	361454	1174
14	459359	5160	360280	1639

Table 26. Data on aggregated forces that are reformatted without air sorties. Forces are combat manpower, APCs, Tanks and artillery weighted by 1, 5, 20 and 40, respectively.

Day	German grnd, att, role sorties	Soviet grnd, att, role sorties
1	160	1
2	1942	600
3	1356	613
4	1499	661
5	1426	669
6	1286	472
7	530	383
8	809	348
9	460	603
10	451	623
11	1147	704
12	541	369
13	278	681
14	122	336
15	18	377

Table 27. Data on number of ground attack role air sorties for German and Soviet forces.

Day	Blue forces	Blue losses	Red forces	Red losses
1	642512	11167	483277	11257
2	631735	12993	454440	9532
3	620182	16266	449198	6249
4	604156	16472	440759	5702
5	581774	18071	430857	6043
6	561033	14445	402134	3450
7	545538	10754	407054	4415
8	542434	28492	392169	5112
9	514542	13302	386787	3491
10	503670	11323	404176	3290
11	482297	6201	382706	3047
12	485456	3600	371769	1975
13	471506	2067	365114	1174
14	470669	5160	360820	1639

Table 28. Data on aggregated forces reformatted with air sorties. Forces are combat manpower, APCs, tanks, artillery and number of ground attack role air sorties which are weighted by 1, 5, 20,40 and 30, respectively.

a. *Estimation of Parameters*

After reformatting the data, linear regression is applied to logarithmically transformed Lanchester equations to estimate the model parameters which are given in equations IV.A.2.(12) and IV.A.2.(13).

To estimate the parameters of the model, which minimize the sum of squared residuals, 14 days of data given in Table 24, Table 26 and S-PLUS Software are used.

b. *Results*

Results for the models are:

Fricker's model for the Kursk data without the air sorties, with tactical parameter *d*, with an SSR value of 5.94×10^{8} and an R^2 value of 0.1703 is:

$$\dot{B} = 3.77\times10^{-33}\left(\frac{100}{79} \text{ or } \frac{79}{100}\right)R^{0.0604}B^{6.3066} \tag{14}$$

$$\dot{R} = 1.09\times10^{-32}\left(\frac{79}{100} \text{ or } \frac{100}{79}\right)B^{0.0604}R^{6.3066} \tag{15}$$

Fricker's model for the Kursk data without the air sorties, and without the tactical parameter *d*, with an SSR value of 6.69×10^{8} and an R^2 value of 0.0657 is:

$$\dot{B} = 1.19\times10^{-32} R^{3.6736} B^{2.6934} \tag{16}$$

$$\dot{R} = 3.44\times10^{-33} B^{3.6736} R^{2.6934} \tag{17}$$

It is significant that the resulting parameters are sensitive to the *d* parameter; after adding the *d* parameter, the *p* and *q* parameters change dramatically.

The above parameters are the ones that give the smallest SSR value. It is possible to have smaller SSR values for the model with the tactical parameter *d* if the parameter *p* or *q* is allowed to have negative values. In staying consistent with Fricker's approach, negative exponent parameters are not considered in this section. Negative values are looked at in the conclusion section.

Fricker's model for the Kursk data with the air sorties, with tactical parameter *d*, with an SSR value of 6.24×10^{8} and an R^2 value of 0.1285 is:

$$\dot{B} = 3.35\times10^{-27}\left(\frac{100}{93} \text{ or } \frac{93}{100}\right) R^{0.0955} B^{5.2207} \tag{18}$$

$$\dot{R} = 5.76\times10^{-27}\left(\frac{93}{100} \text{ or } \frac{100}{93}\right) B^{0.0955} R^{5.2207} \tag{19}$$

Fricker's model for the Kursk data with the air sorties, and without the tactical parameter *d*, with an SSR value of 7.18×10^{8} and an R^2 value of –0.020 is:

$$\dot{B} = 5.01\times10^{-27} R^{1.4983} B^{3.8179} \tag{20}$$

$$\dot{R} = 3.85\times10^{-27} R^{1.4983} B^{3.8179} \tag{21}$$

Like the models without the air sorties added, the above parameters are the ones that give the smallest SSR value. It is possible to have smaller SSR values for the model with the tactical parameter *d* if the parameter *p* or *q* is allowed to have negative

values. That is the algorithm of a force's casualties decreases as one of the force strengths increases, and since this interpretation does not make sense, the negative values are not considered.

Figures 22 and 23 show the fitted losses plotted versus real losses for the Soviet and the German forces, respectively, for Fricker's model without the air sortie data added and using the *d* parameter.

Figures 24 and 25 show fitted losses plotted versus real losses for the Soviet and the German forces, respectively, for Fricker's model without the air sortie data added and without using the *d* parameter.

Figures 26 and 27 show the fitted losses plotted versus real losses for the Soviet and the German forces, respectively, for Fricker's model with the air sortie data added and using the *d* parameter.

Figures 28 and 29 show the fitted losses plotted versus real losses for the Soviet and the German forces, respectively, for Fricker's model with the air sortie data added and without using the *d* parameter.

When the R^2 values above, which are found by using Fricker's methodology, are compared, it is seen that adding the air sortie data improves the fit for the Battle of Kursk data. Using the tactical parameter does not improve the fit to the Battle of Kursk data for the model without the air sorties. On the contrary, using the tactical parameter improves the fit to the Battle of Kursk data for the model with the air sorties.

The *d* parameter is found to be 0.79 and 0.93 for the models without the air sorties and with the air sorties, consecutively. This result implies a defender advantage/attacker

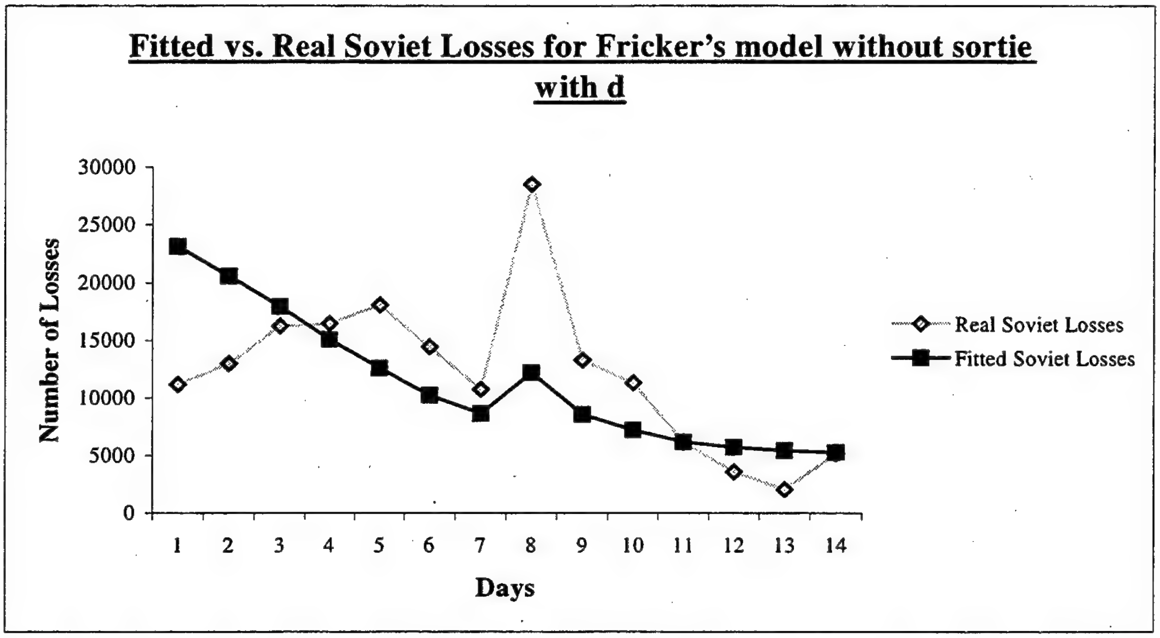

Figure 22. Fitted losses plotted versus real losses for the Soviet forces for Fricker's model without the air sortie data added and using the *d* parameter. Notice the pattern where the model overestimates the initial and the last part of the battle, while underestimating the part in between.

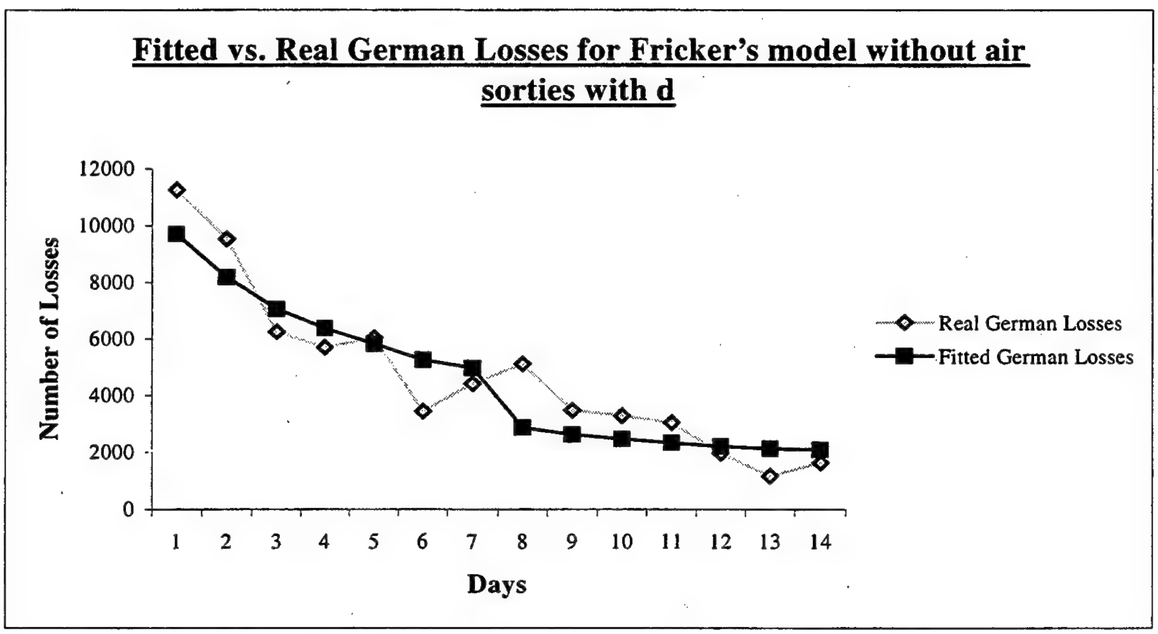

Figure 23. Fitted losses plotted versus real forces for the German forces for Fricker's model without the air sortie data added and using the *d* parameter.

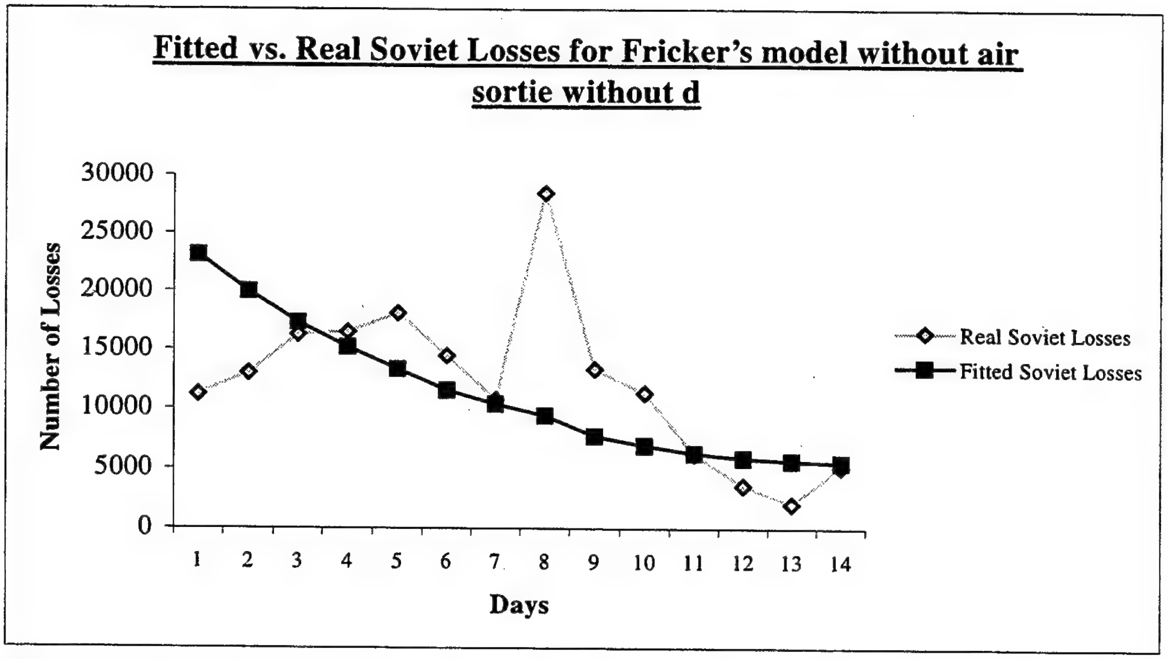

Figure 24. Fitted losses plotted versus real Soviet losses for the Soviet forces for Fricker's model without the air sortie data added and without using the *d* parameter. The same pattern in which the model over/underestimates the battle in three distinctive phases is also observable in this plot.

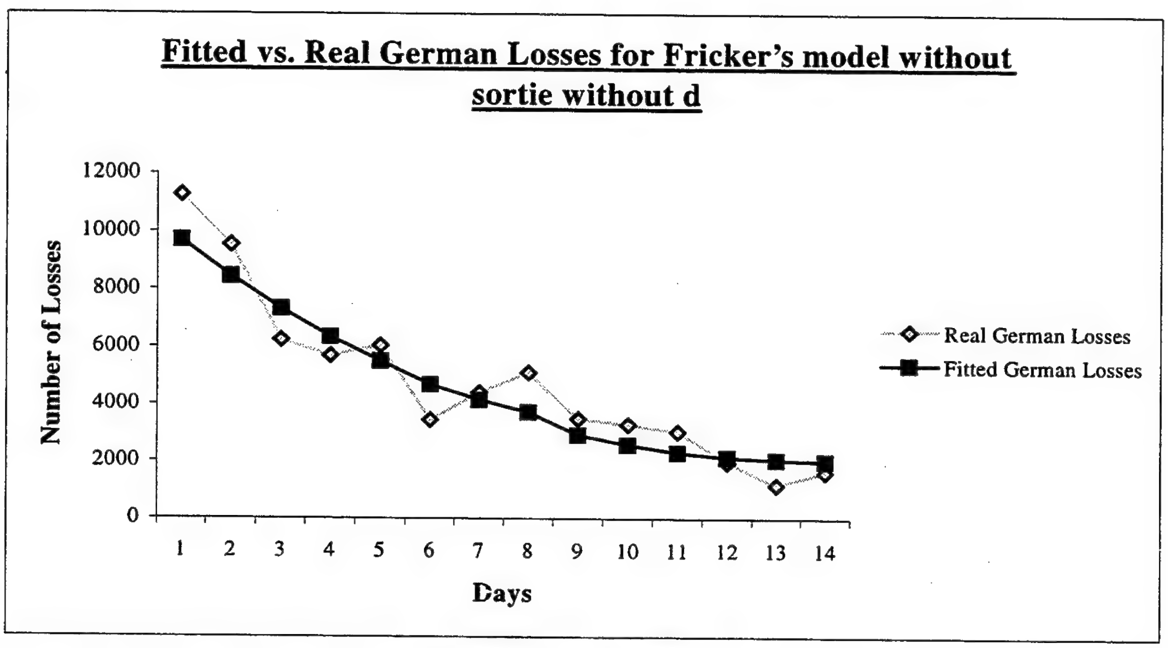

Figure 25. Fitted losses plotted versus real Geman losses for the German forces for Fricker's model without the air sortie data added and without using the *d* parameter.

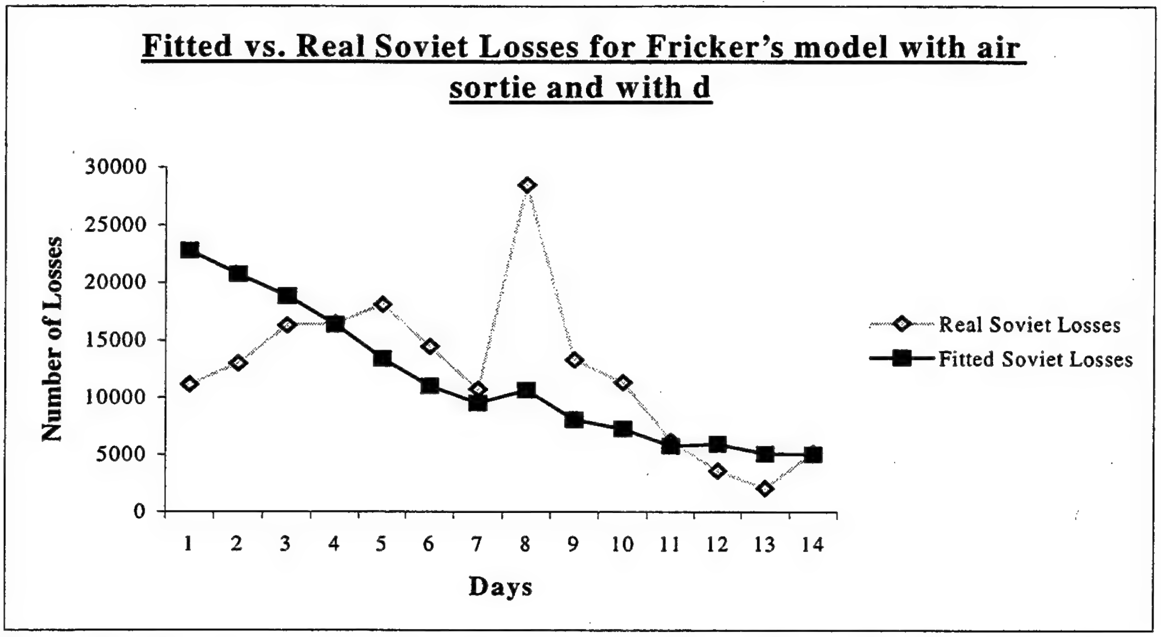

Figure 26. Fitted losses plotted versus real losses for Soviet forces for Fricker's model with the air sortie data added and using the *d* parameter.

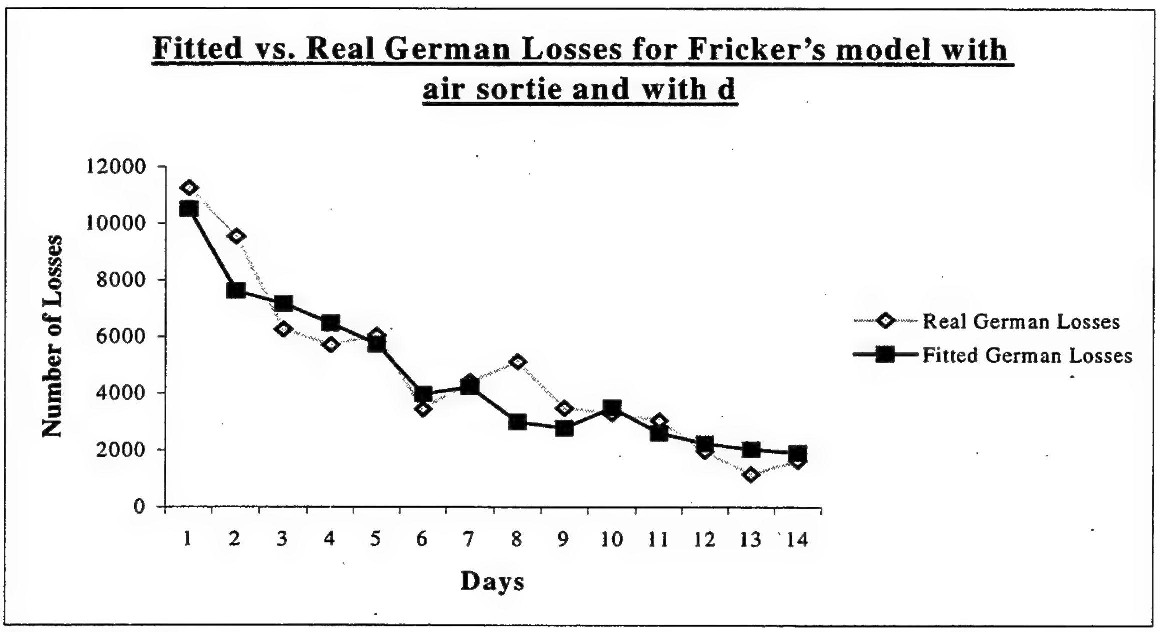

Figure 27. Fitted losses plotted versus real losses for German forces for Fricker's model with the air sortie data added and using the *d* parameter.

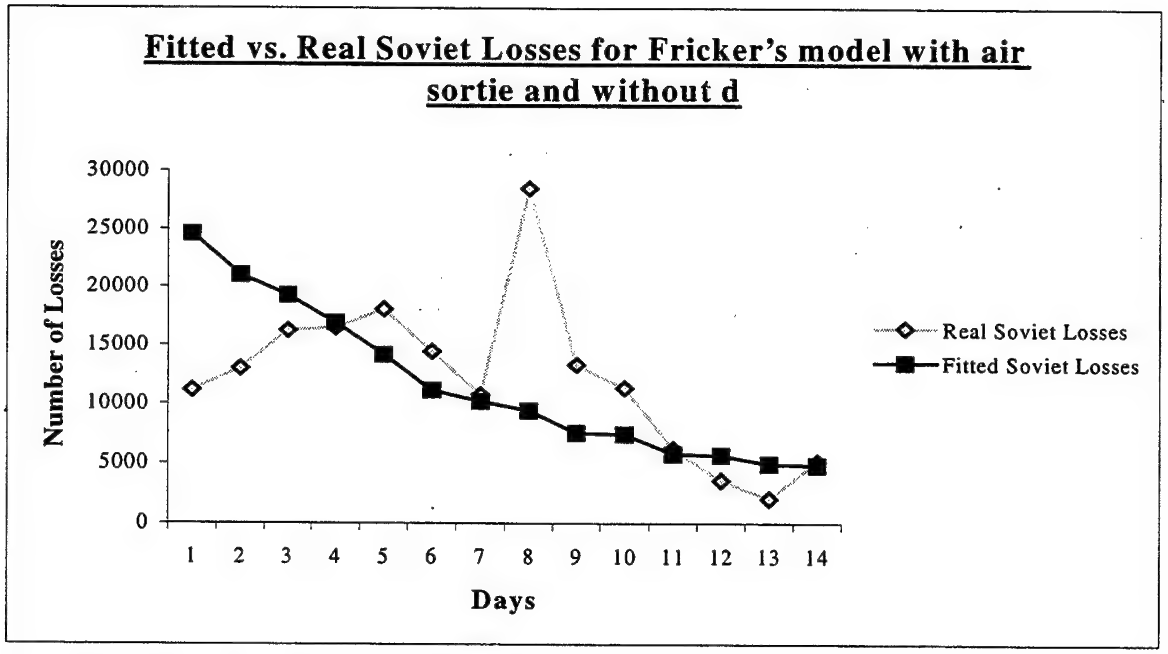

Figure 28. Fitted losses plotted versus real losses for the Soviet forces for Fricker's model with the air sortie data added and without using the *d* parameter.

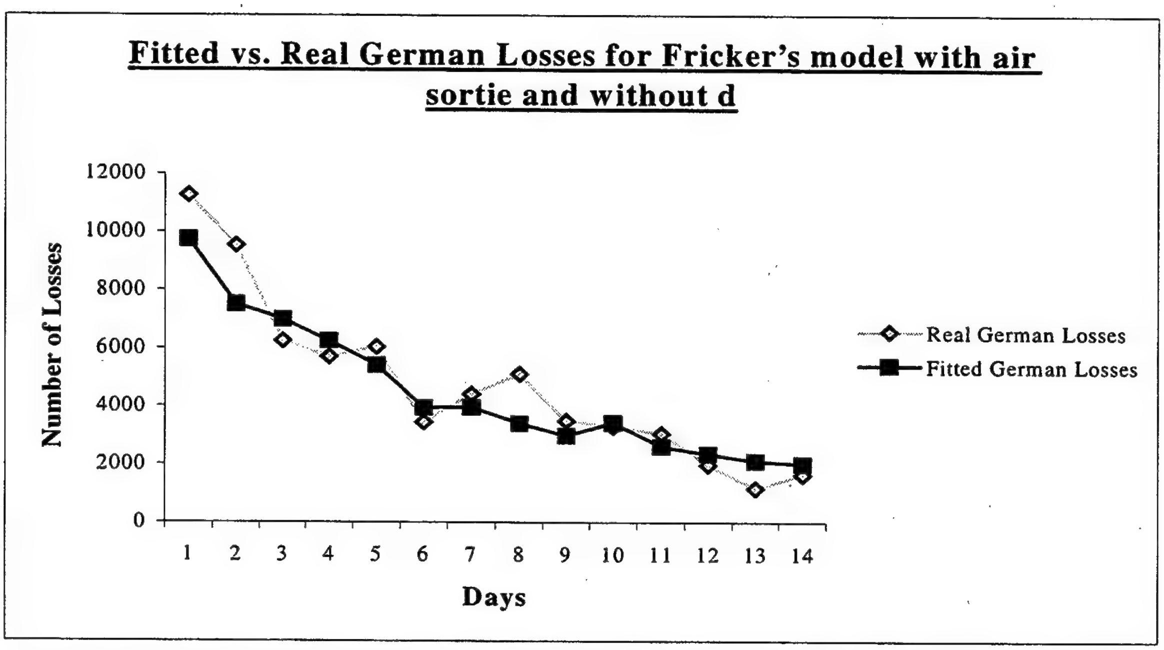

Figure 29. Fitted losses plotted versus real losses for the German forces for Fricker's model with the air sortie data added and without using the *d* parameter.

disadvantage. In both cases, when the tactical parameter is used, $a<b$, and when tactical parameter is not used, $a>b$. Again, in both cases, the a and b parameters are very small.

When the plots given in Figures 22, 24, 26 and 28 are examined, there appears to be three phases in the battle. It is also apparent that the battle lost its intensity after July 12. Notice the pattern where the model overestimates the beginning part and the last part of the battle while underestimating the 8 days in a row between these two parts. This pattern suggests that fitting a model with change points may improve the model's fit to the data. Also, the model provides a much better fit for the German side.

In equations IV.A.2.b.(14), IV.A.2.b.(15), IV.A.2.b.(18), IV.A.2.b.(19) and IV.A.2.b.(20), IV.A.2.b.(21) the q parameter is greater than the p parameter suggesting that one side's loss is more a function of his own forces rather than his opponent's forces. This finding is similar to what Fricker observed in his study.

In equations IV.A.2.b.(16), IV.A.2.b.(17) the p parameter is greater than the q parameter, which suggest that one side's loss is more a function of his opponent's forces rather than his own forces. This finding is different from Fricker's findings.

It is significant that using the tactical parameter d does not improve the fit for the model without the air sortie data when SSR values are compared. This may be interpreted as using the logarithmically transformed equations does not necessarily gives the best fit in the original form. Table 29 shows the results for Fricker's models as a whole for both the Ardennes and the Kursk data. The negative R^2 values found here imply that the fitted model yields worse results than using the average daily losses as an estimate. This finding was communicated with Fricker and it was concluded that the reason for the negative R^2 values are the combination of extreme sensitivity of the

results to the precision of parameters and using the rounded off values given in Fricker's study [Ref.6]. For example, for the first model given in Table 29, changing the q parameter from 5.0 to 5.02 increases the R^2 value from –0.7938 to 0.1904, and changing the q parameter from 5.0 to 5.03 increases the R^2 value to 0.4581.

Name of the model	a	b	p	q	d	SSR	R^2
Ardennes w/o sorties with d	4.7E-27	3.1E-26	0	5	0.8093	1.57E+8	-0.7938
Ardennes w sorties with d	2.7E-24	1.6E-23	0	4.6	0.7971	2.64E+7	0.5256
Kursk w/o sorties with d	3.76E-33	1.09E-32	0.0604	6.3066	0.79	5.94E+8	0.1703
Kursk w/o sorties w/o d	1.61E-33	3.44E-33	3.6736	2.6934	-	2.16E+9	0.0657
Kursk with sorties with d	3.35E-27	5.76E-27	0.0955	5.2207	0.93	6.23E+8	0.1294
Kursk with sorties w/o d	5.01E-27	3.85E-27	1.4983	3.8179	-	7.16E+8	-0.0222

Table 29. Fricker's results for his models with/without the tactical parameter d, with/without the air sortie added, for both the Ardennes and the Kursk data.

B. EXPLORATORY ANALYSIS OF BATTLE OF KURSK DATA

The fighting on the first day of the battle was sporadic. The extremely low casualty levels represent large outliers; this, including the data of the first day could drastically effect the outcome of the analysis. Thus, the first day of data was dropped in fitting the data to the models. This kind of approach is also supported by the historical account of the Battle of Kursk, because the main offensive did not really begin until July 5, the second day of the battle. Even if there are other days on which large outliers are observed—like July 12—these outliers will not be left out of the analysis as they are a

result of the fighting during the Kursk offensive. Therefore, this study will fit only the last 14 days of the aggregated data given in Table 14, excluding the first day. All the results found from the models are summarized as a whole in Table 42 in Section IV.B.10.

1. The scalar aggregation models

Two numerical methods are used to fit parameters to the scalar model of Lanchester equations. One is linear regression and the other is robust LTS regression. Robust LTS regression method performs least-trimmed squares regression [Ref.17]. When the given data in hand contains significant outliers as in our case, robust regression models are useful for fitting linear relationships by discounting outlying data. Both methods minimize the sum of squared residual (SSR) error resulting from the model to the actual data.

a. Linear regression

Linear regression is used for fitting parameters to the logarithmically transformed Lanchester equations. The original form of Lanchester equations are given in equations I.A.(1) and I.A.(2). By taking the logarithm of each side of the equations, we get:

$$\log(\dot{B}) = \log(a) + p\log(R) + q\log(B) \tag{22}$$

$$\log(\dot{R}) = \log(b) + p\log(B) + q\log(R) \tag{23}$$

Only the last 14 days of the data given in Table 19 are used for performing the linear regression analysis.

b. Results of the linear regression model

Results of the linear regression model which gives an SSR value of 6.36×10^{8} and an R^2 value of 0.1126 are:

$$\dot{B} = 1.06\times10^{-47} R^{5.7475} B^{3.3356} \tag{24}$$

$$\dot{R} = 1.90\times10^{-48} B^{5.7475} R^{3.3356} \tag{25}$$

c. ***Robust LTS regression***

We will use Robust LTS regression for fitting parameters to the Lanchester equations. The original form of Lanchester equations are given in equations I.A.(1) and I.A.(2). By taking the log of each side of the equations, we obtain the equations given in IV.A.2.(22) and IV.A.2(23). Only the last 14 days of the data given in Table 19 are used for doing the robust LTS regression analysis.

d. ***Results of the robust LTS regression***

Results for the robust LTS regression model which gives an SSR value of 5.54×10^{8} and an R^2 value of 0.2262 are:

$$\dot{B} = 2.27\times10^{-40} R^{6.0843} B^{1.7312} \tag{26}$$

$$\dot{R} = 1.84\times10^{-41} B^{6.0843} R^{1.7312} \tag{27}$$

Figures 30 and 31 show fitted losses plotted versus real losses for the Soviet and the German forces, respectively, for the linear regression model. Figures 31 and 32 show fitted losses plotted versus real losses for the Soviet and the German forces, respectively, for the robust LTS regression model.

When the SSR values found by using linear regression and robust LTS regression techniques are compared, it is observed that using the robust LTS regression technique improves the fit for the Battle of Kursk data. The SSR value, which is found by using the robust LTS regression method, is the smallest for the Kursk data so far.

It should be noted that even if the robust LTS regression technique accounts for the outliers when finding the parameters that minimize the SSR for a given model, the

SSR values computed here include the SSR of the outliers. In other words, when the parameters computed by the robust LTS regression technique are used in the analysis, the

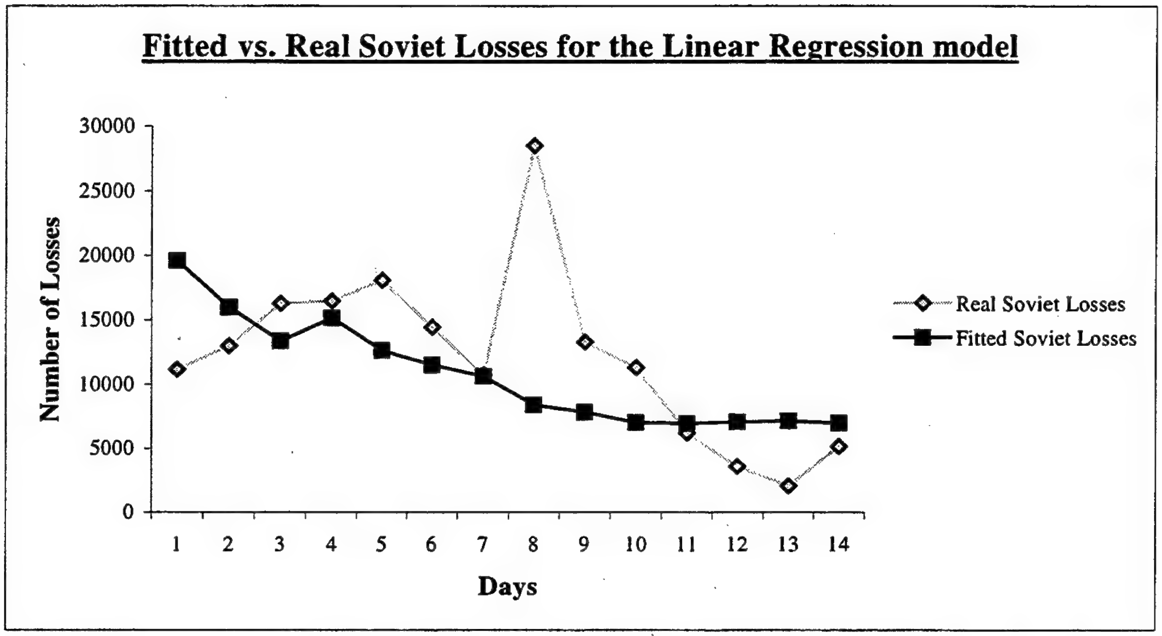

Figure 30. Fitted losses plotted versus real losses for Soviet forces for the linear regression model. The significant outlier on day 8 influences the fit dramatically.

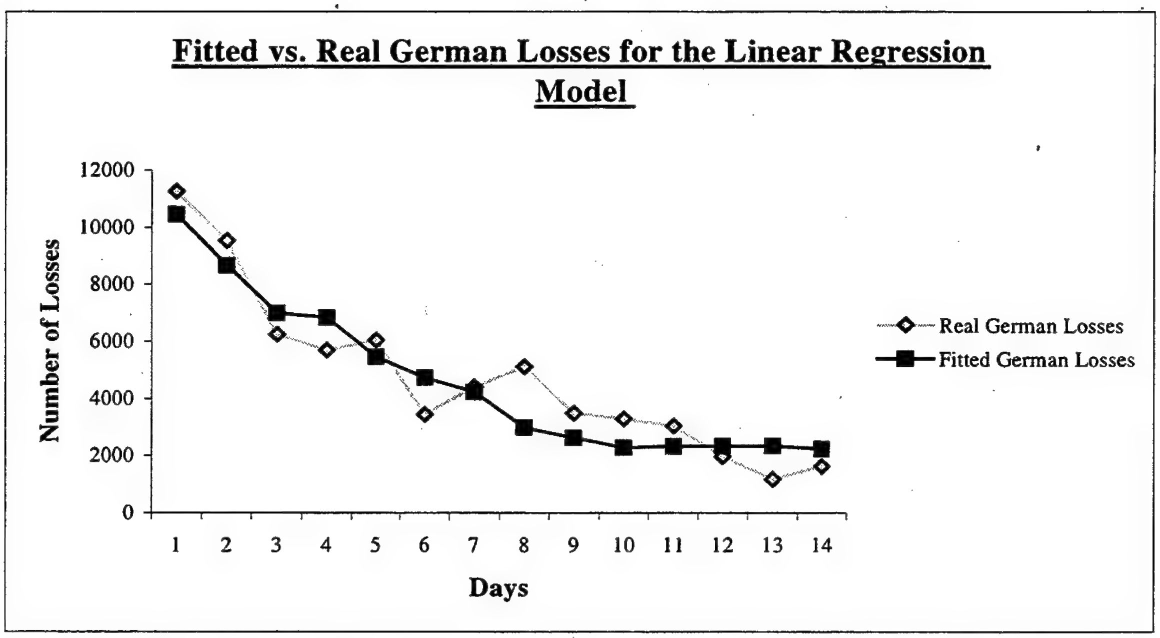

Figure 31. Fitted losses plotted versus real losses for German forces for the robust LTS regression model. The data for the German side, with no significant outliers, gives a better fit for the model.

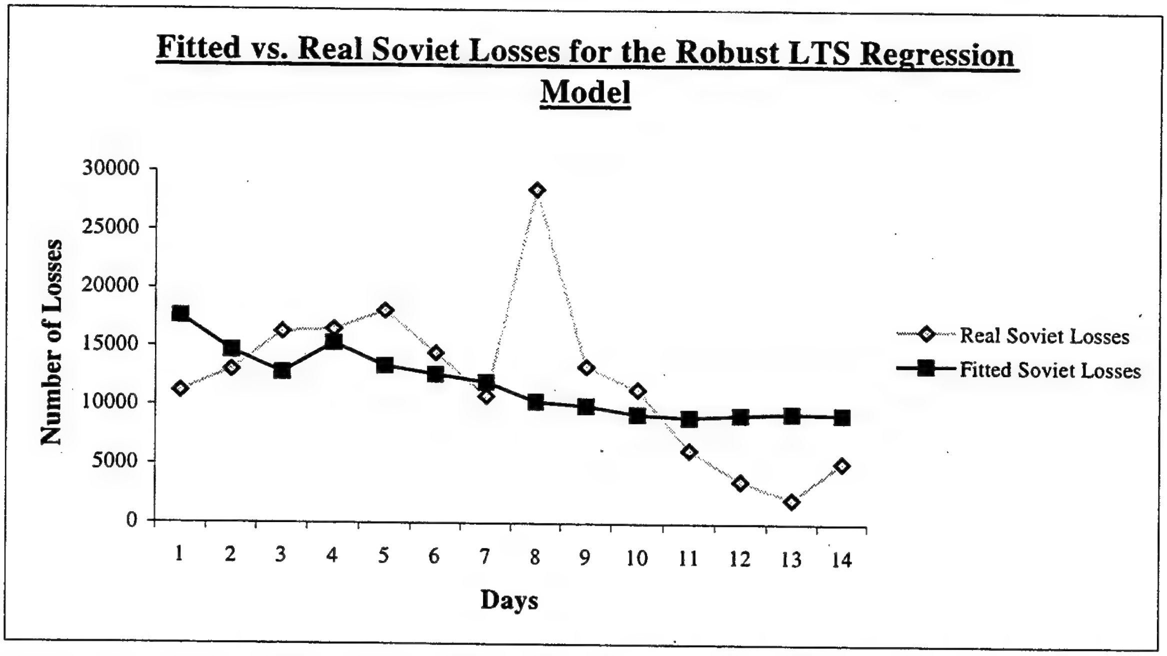

Figure 32. Fitted losses plotted versus real losses for Soviet forces for the robust LTS regression model. The significant outlier on day 8 influences the fit dramatically.

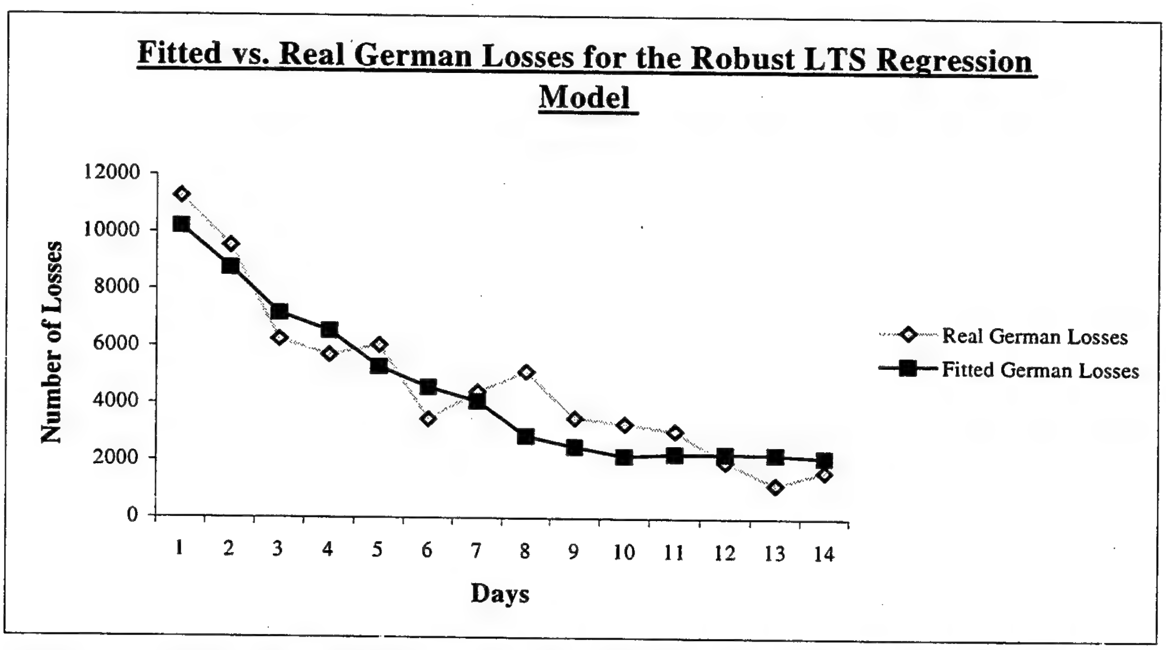

Figure 33. Fitted losses plotted versus real losses for the German forces for the robust LTS regression model. The data for the German side, with no significant outliers, gives a better fit for the model.

outliers are not discounted. They are included for the purpose of computing the SSR value.

When the p and q parameters are compared it is noticed that the p parameter is greater than the q parameter, suggesting that one side's loss is more a function of his opponent's forces rather than his own forces. This interpretation is different from what Fricker found in his study. In both cases, a and b parameters are significantly small, and $a > b$.

When the plots given in Figures 30 and 32 are examined, there appears to be three distinct phases in the battle. It is also apparent that the battle lost its intensity after July 12. After the Soviets went into offense, the battle was not as intense. There is a clear pattern in Figure 30 where the model overestimates the beginning part and the last part of the battle, while underestimating the attrition for eight days in a row between these two periods.

The pattern seen in Figures 30 and 32 suggests that fitting a model with change points may improve the model's fit to the data. Likewise, leaving out the data given for July 12 when the most intense fighting of the battle took place, it may also be possible to increase the fit to the data, an approach which will be covered in upcoming sections. Also, the model provides a much better fit for the German side.

2. Including air sortie data

As mentioned in IV.A.2, the air sortie data given in the KOSAVE study [Ref.12] consists of the number of air-air role sorties, ground attack role sorties, reconnaissance role sorties and evacuation role sorties (which are solely used by Germans). For aggregating the air sortie data into total aggregated number of forces, we will use the data

given in Table 27 that presents data on the number of ground attack role sorties. However, the aggregated data will be different than that given in Table 28, because the data in Table 28 is calculated using the reformatted data by applying Fricker's algorithm.

The data, which we will be using in this section, is given in Table 30 which presents the total number of aggregated forces, including the air data by weighing each sortie by 30. In other words, the number of air sorties presented in Table 27 is multiplied by 30 and added onto the aggregated force levels given in Table 19 in order to compute the data presented in Table 30.

Two regression methods, presented in IV.B.1 are used for fitting the data given in Table 30, namely, linear regression and robust LTS regression.

day	Blue Forces	Blue Losses	Red Forces	Red Losses
1	604353	11167	431671	11257
2	594159	12993	404945	9532
3	579175	16266	404055	6249
4	565402	16472	415304	5702
5	542712	18071	406024	6043
6	527893	14445	382404	3450
7	518016	10754	389340	4415
8	498123	28492	375765	5112
9	487961	13302	375759	3491
10	480724	11323	394230	3290
11	474229	6201	373752	3047
12	482881	3600	367286	1975
13	471266	2067	363905	1174
14	469253	5160	360820	1639

Table 30. Data on aggregated forces. Forces are combat manpower, APCs, tanks, artillery and number of ground-attack role sorties which are weighted by 1, 5, 20, 40 and 30, respectively.

a. Results of linear regression model

Results for the linear regression model, which gives an SSR value of 6.85×10^8 and an R^2 value of 0.0433, are:

$$\dot{B} = 1.40\times10^{-35} R^{5.1323} B^{1.7793} \tag{28}$$

$$\dot{R} = 2.09\times10^{-36} B^{5.1323} R^{1.7793} \tag{29}$$

b. Results of robust LTS regression model

Results for the robust LTS regression model, which gives an SSR value of 7.58×10^{8} and an R^2 value of -0.0579, are:

$$\dot{B} = 1.21\times10^{-38} R^{5.3691} B^{2.0883} \tag{30}$$

$$\dot{R} = 1.75\times10^{-39} B^{5.3691} R^{2.0883} \tag{31}$$

Figures 34 and 35 show the fitted losses plotted versus real losses for the Soviet and the German forces respectively, for the linear regression model with the air sortie data added.

Figures 36 and 37 show the fitted losses plotted versus real losses for the Soviet and the German forces respectively, for the robust LTS regression model with the air sortie data added.

Following the aggregation of the data using the number of air sorties, it is not appropriate to compare the models using the SSR values because, the increase in the SSR value may be a natural result of adding the air sortie data. For this reason, R^2 values will be used to compare the fit of the model.

Upon the examination of the R^2 values above, which are found by applying linear regression and robust LTS regression techniques to the logarithmically transformed data that includes air sorties, one can determine that considering the air sorties data does not improve the model's fit to the data. The R^2 values, which are found by using the linear regression and the robust LTS regression technique, are both lower than the R^2 values

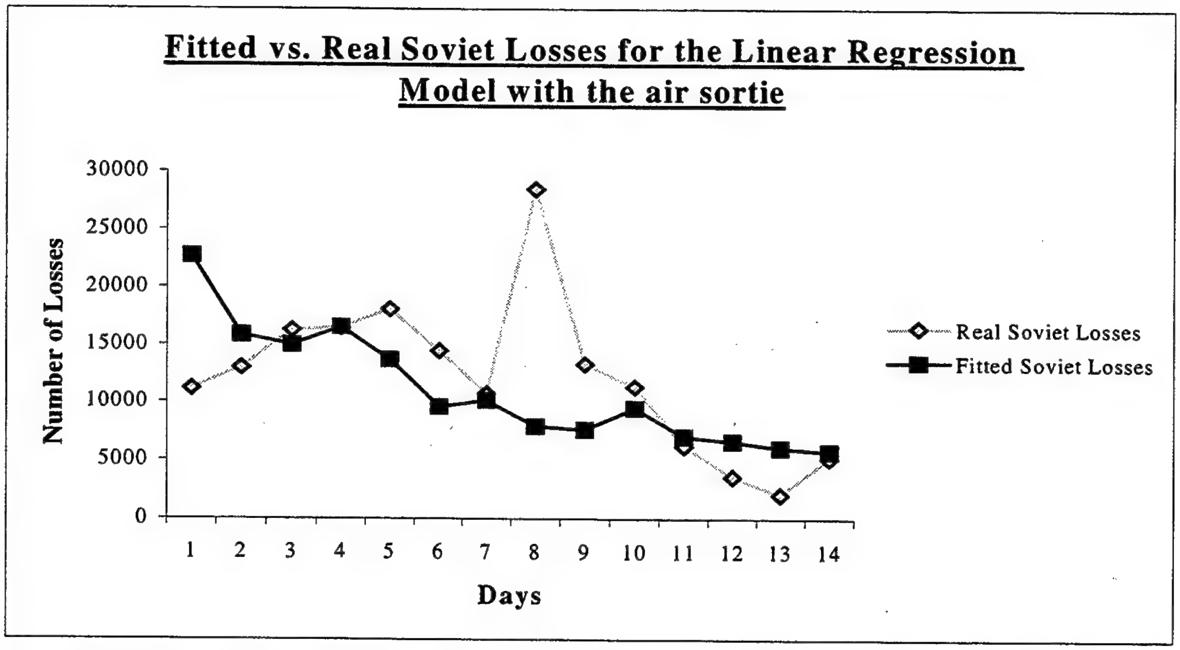

Figure 34. Fitted losses plotted versus real losses for Soviet forces for the linear regression model with the air sortie data added. The significant outlier on day 8 influences the fit dramatically. The same pattern where the model over/underestimates the battle in three distinctive phases is observable in this plot too.

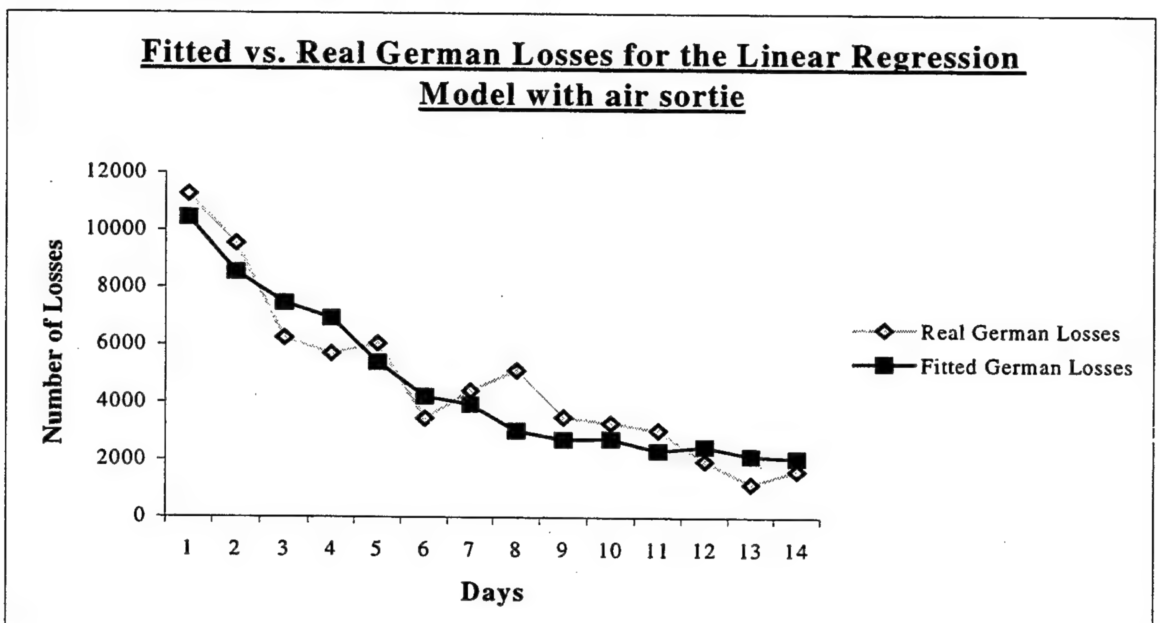

Figure 35. Fitted losses plotted versus real losses for German forces for the linear regression model with the air sortie data added. The data for the German side, with no significant outliers, gives a better fit for the model.

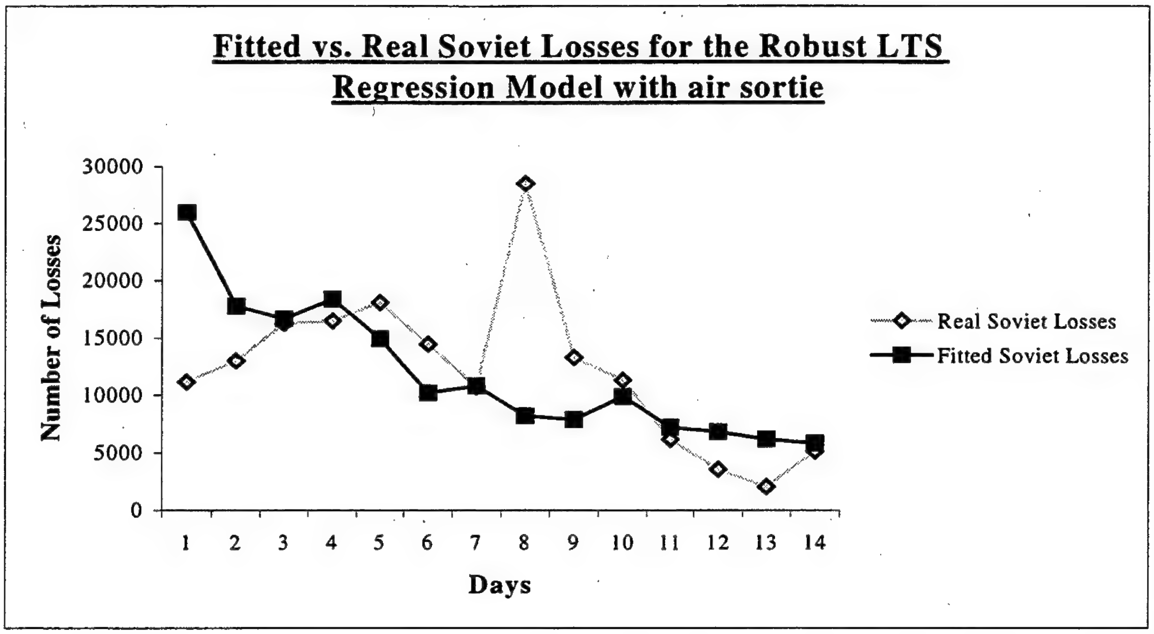

Figure 36. Fitted losses plotted versus real losses for the Soviet forces for the robust LTS regression model with the air sortie data added. The significant outlier on day 8 influences the fit dramatically. The same pattern where the model over/underestimates the battle in three distinctive phases is observable in this plot too.

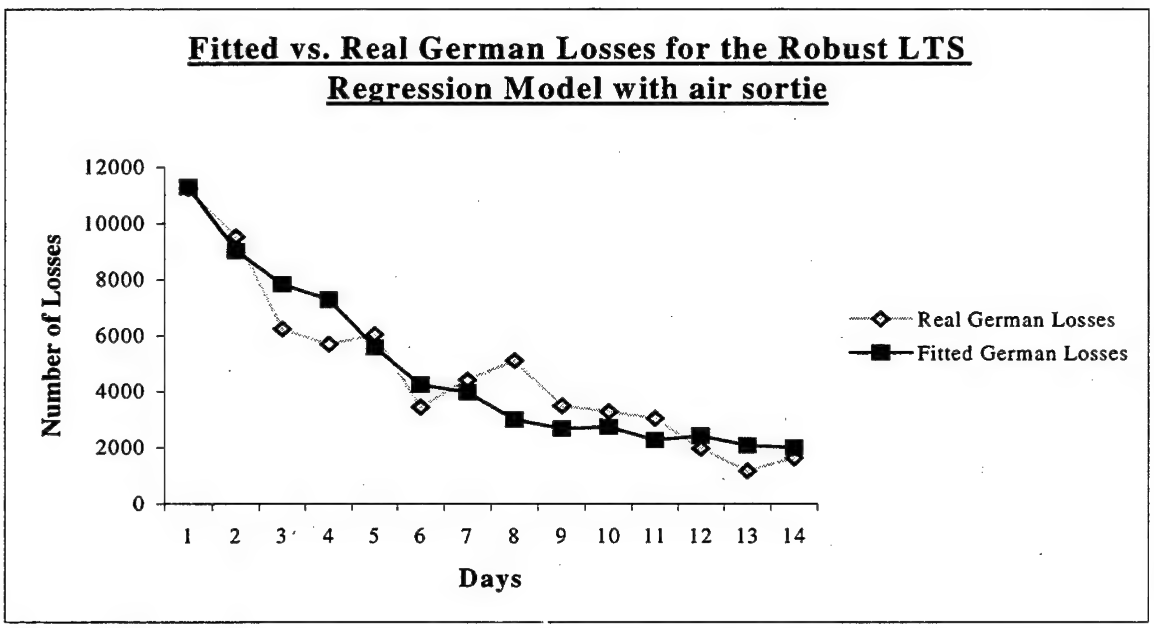

Figure 37. Fitted losses plotted versus real losses for the German forces for the robust LTS regression model with the air sortie data added The data for the German side, with no significant outliers, gives a better fit for the model.

found in Section IV.B.1 which did not include the air sortie data. While the previous best fit found was 0.2262 in section IV.B.1.d, after the air sortie data is added, R^2 is found to be 0.0433 and -0.0579. Adding air sortie data did not improve model's fit to the data.

For both cases, the a and b parameters are significantly small and $a > b$. This suggests that individual German effectiveness was greater than individual Russian effectiveness.

When the p and q parameters are compared it is observed that the p parameter is greater than the q parameter, indicating that one side's losses are more a function of his opponent's forces rather than being a function of his own forces. This result is different from what Fricker found in his study.

When the plots given in Figures 34 and 36 are examined, the resulting pattern is similar to the one seen in the previous section. This pattern again suggests that fitting a model with change points may improve the model's fit to the data. Again, similar to the previous results, it may be possible to increase the fit to the data by leaving out the data given for July 12.

3. Taking into account the change in offensive/defensive roles

By historical account, the German forces generally maintained an offensive posture (this is not valid for all units on the battlefield) through July 12, when the Soviets were able to gain the initiative and launch their counter-offensive. Bracken [Ref.13] introduced an additional parameter d to the standard Lanchester equations (I.B.(1) and I.B.(2)), called a *tactical parameter*, to account for a battle in which defense and offense switch during the course of the campaign.

With d for the defender and *(1/d)* for the attacker, the Lanchester equations are modified to accept the tactical parameter d and are given as:

$$\dot{B} = (d \text{ or } 1/d)\, aR^p B^q \tag{32}$$

$$\dot{R} = (1/d \text{ or } d)\, bB^p R^q \tag{33}$$

The logarithmically transformed Lanchester equations which are modified to accept the tactical parameter (for the days that red is the attacker), are given as:

$$\log(\dot{B}/d) = \log(a) + p\log(R) + q\log(B) \tag{34}$$

$$\log(\dot{R}/(1/d)) = \log(b) + p\log(B) + q\log(R) \tag{35}$$

Linear regression and robust LTS regression models are used to estimate the model parameters represented above in IV.B.3.(34) and IV.B.3.(35).

a. Linear regression

The last 14 days of the aggregated data given in Table 14 in section IV.A.1 and the S-PLUS software are used to estimate the model's parameters, which minimize the sum of squared residuals of the actual and estimated attrition.

In order to iterate for different d values, linear regression is fit for multiple d values, and then the d value that gives the minimum SSR is selected. The value of tactical parameter d is varied between 0.0 and 9.0 in increments of 0.01.

b. Results of the linear regression model

Results for the linear regression model which gives an SSR value of 6.24×10^8 and a tactical parameter value of 1.17 and an R^2 value of 0.1295 are:

$$\dot{B} = (\frac{1}{1.17} \text{ or } 1.17)\, 1.88\times10^{-47} R^{7.5038} B^{1.5793} \tag{36}$$

$$\dot{R} = (1.17 \text{ or } \frac{1}{1.17})\, 1.07\times10^{-48} B^{7.5038} R^{1.5793} \tag{37}$$

c. *Robust LTS regression*

For estimating the parameters, which minimize the sum of squared residuals of the actual and estimated attirition, the last 14 days of the aggregated data given in Table 5, in Section IV.A.1 and the S-PLUS software are used.

d. *Results of the robust LTS regression*

Results for the robust LTS regression model which gives an SSR value of 5.54×10^8 and a tactical parameter value of 1.00 and an R^2 value of 0.2262 are:

$$\dot{B} = (\frac{1}{1.0} \text{ or } 1.0) 2.27 \times 10^{-40} R^{6.0843} B^{1.7312} \tag{38}$$

$$\dot{R} = (1.0 \text{ or } \frac{1}{1.0}) 1.84 \times 10^{-41} R^{6.0843} B^{1.7312} \tag{39}$$

Figures 38 and 39 shows the fitted losses plotted versus real losses of the Soviet and the German forces, respectively, for the linear regression model.

Figures 40 and 41 shows the fitted losses plotted versus real losses of the Soviet and the German forces, respectively, for the robust LTS regression model.

When the SSR values above are examined, it is apparent that taking into consideration the change in offensive/defensive roles improves the fit. The SSR values, which are found by using the linear regression and robust LTS regression technique, are both less than or equal to the SSR values found in section IV.B.1, which did not consider the change in offensive/defensive roles. The best fit found in section IV.B.1 was 6.36×10^8 for the linear regression model, after the d parameter is included in the model, SSR value is found to be 6.24×10^8, suggesting only a 2% improvement in fit. But, this is not the case for robust LTS regression model. While the previous result for robust LTS regression model was found to be 5.54×10^8 in Section IV.B.1.d, after the change in

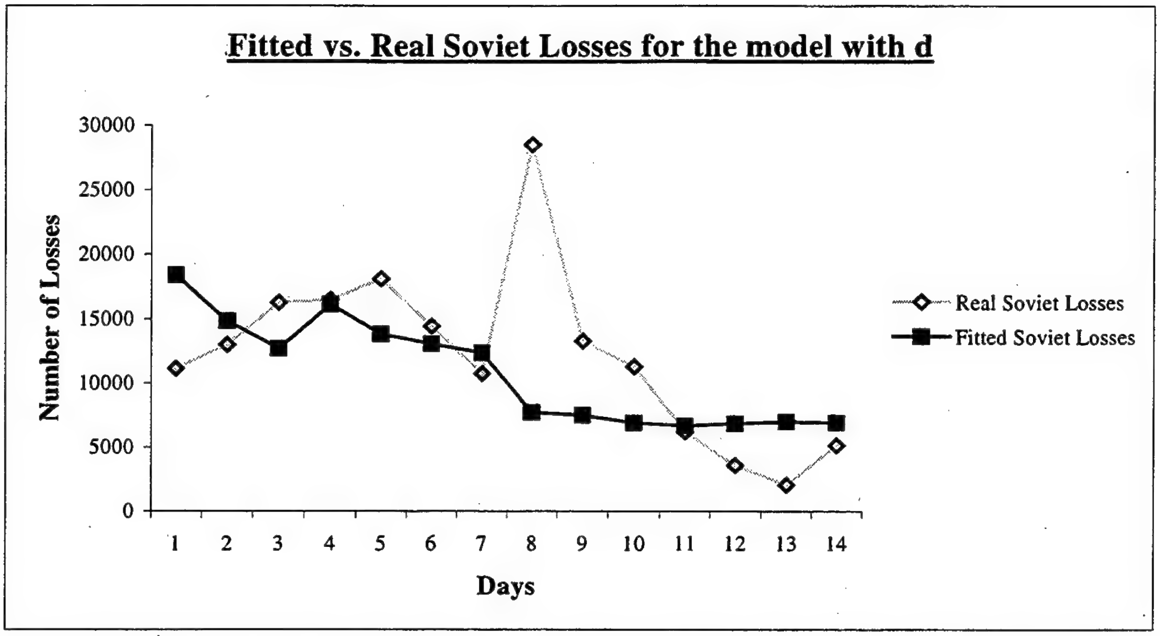

Figure 38. Fitted losses plotted versus real losses for the Soviet forces for the linear regression model with the tactical parameter *d*. The significant outlier on day 8 influences the fit dramatically.

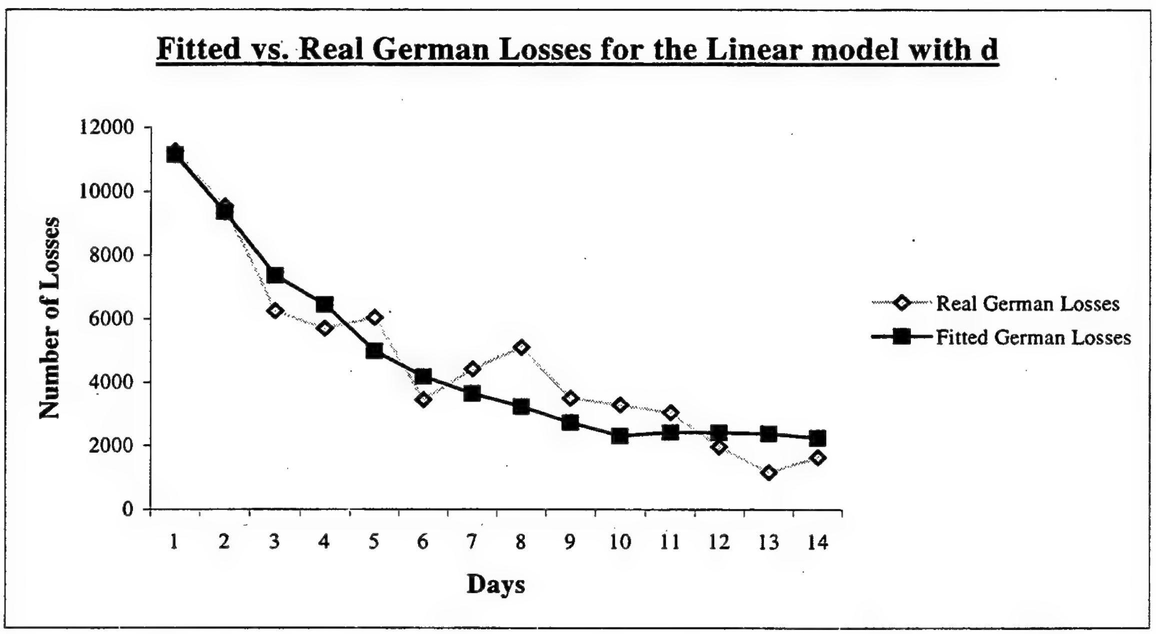

Figure 39. Fitted losses plotted versus real losses for the German forces for the linear regression model with the tactical parameter *d*. The data for the German side, with no significant outliers gives a better fit for the model.

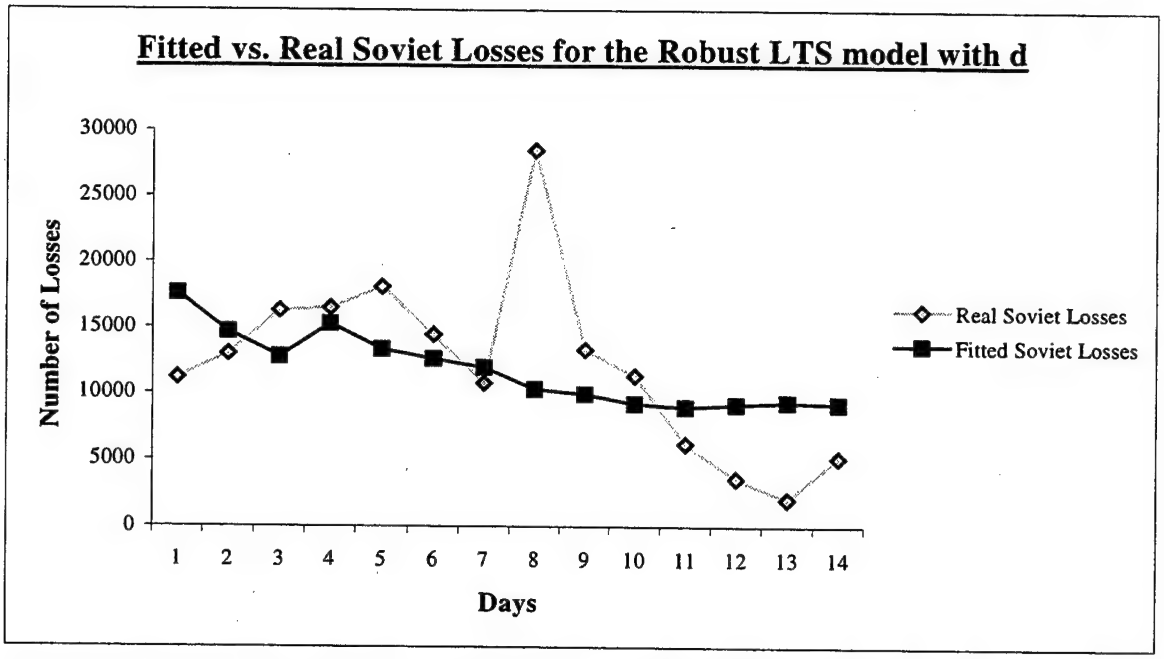

Figure 40. Fitted losses plotted versus Real losses for the Soviet forces for the robust LTS model with the tactical parameter *d*.

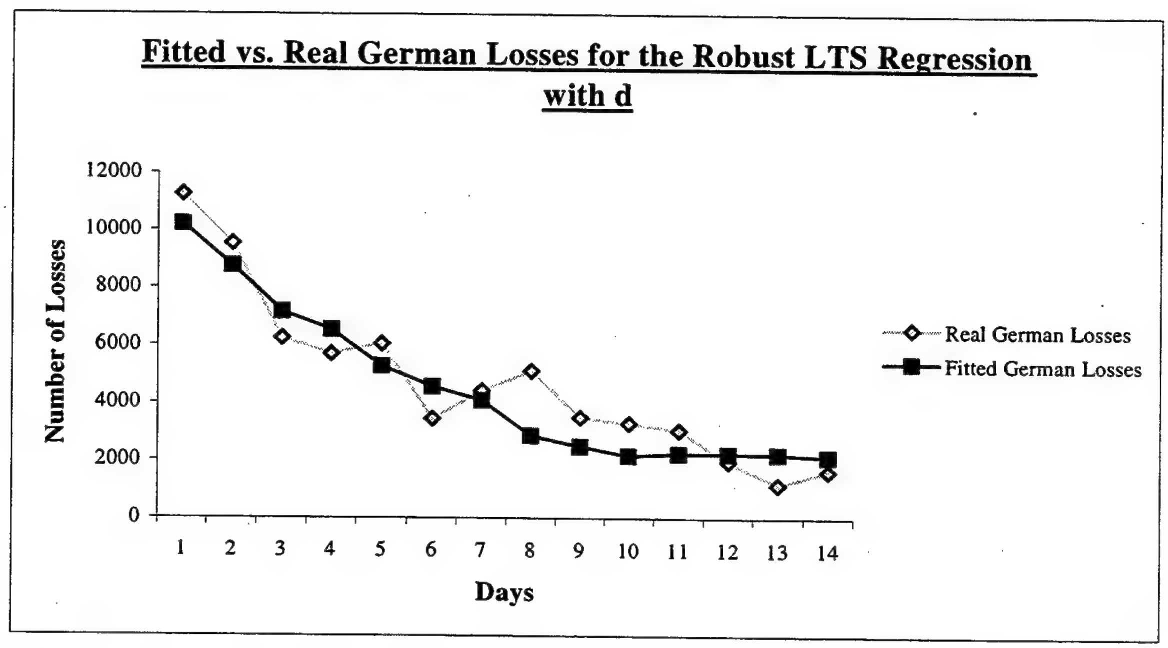

Figure 41. Fitted losses plotted versus Real losses for the German forces for the robust LTS model with the tactical parameter *d*.

offensive/defensive roles is taken into account, it is again found to be 5.54×10^{8}. In other words, taking into account the change in offensive/defensive roles does not change the fit for the robust LTS regression model.

Following a search of the tactical parameter *d* value, performed in increments of 0.01, 1.0 is found to be the optimal *d* value that gives the smallest SSR value for the robust LTS regression model. This result indicates that in the context of the Battle of Kursk, one side's status as the defender or attacker does not affect the number of losses which either of the sides is going to suffer. This reasoning may not intuitively make sense, but further analysis made in the following sections will provide additional rationale.

For both cases, the *a* and *b* parameters are significantly small, and $a>b$. This suggests that individual German effectiveness is greater than individual Russian effectiveness.

The *d* parameter with a value of 1.17 signifies that the attacker has an advantage. This result is somewhat unexpected and implies that it is the attacker who will suffer fewer casualties. (The *d* parameter is investigated more closely in upcoming sections).

When *p* and *q* parameters are compared, it is observed that the *p* parameter is greater than the *q* parameter, suggesting that one side's losses are more a function of the opponent's forces rather than a function of its own forces. This finding is different from what Fricker found in his study.

When the plots given in Figures 38 through 41 are examined, the pattern seen in these plots are similar to the results observed in the previous section. This pattern, again,

suggests that fitting a model with change points may improve the fit and again the models fit better for the Germans.

4. Considering the tactical parameter *d* of the campaign

The findings in the previous sections suggest that fitting models with different d values for separate phases of the battle might improve the fit to the data and this section focuses on that aspect of our findings and will analyze the battle in separate time periods.

The tactical parameter found in the previous section, d=1.17, is similar to Bracken's [Ref.8] findings which also implied an attacker advantage. Since $d>1$, implying that if Blue is defending, then blue has a defender disadvantage, and if red is attacking when $d>1$, then red has an attacker advantage. This intuitively does not make much sense because the defender is usually dug in, and the attacker is out in the open and easily detected by the enemy. It should be the defender who has the advantage rather than the attacker when attrition rates are considered. In this situation, it may not make sense to have only one d for the whole campaign.

A closer look at the battle data may find a better fit for the model. The very first day of the battle, the Germans run into the heavily fortified Soviet positions and minefields and have a very rough day. This first day, the Germans obviously have an attacker disadvantage, while the Russians have a defender advantage. July 6, 1943 is the day when things begin to run smoothly for the Germans, as they are not up against a fortified defense, dense barriers and minefields. This scenario continues until July 12, when the Soviets launch their counter-attack. Even on that day, the Germans were not aware of the Soviets' intention to make such a move [Ref.16]. July 12, 1943 can be viewed as the day, when neither side was a defender. Both sides attacked each other

resulting in the bloodiest day of the campaign. The Soviets especially suffered heavy casualties. From July 13 on, the Soviets continued their counter-attacks until they recaptured the ground they had lost. During this time Germans use a hasty defense.

This type of approach is also justified by the historical account of the battle, which is explained in detail in [Ref.15] and [Ref.16]. As a result of the clearly defined phases of the battle, the data will be handled in four different time periods. A different d value will be used for each part of the campaign (i.e. there will be four different d parameters for the campaign). A weakness of this approach is the fact that it requires fitting 8 parameters with 14 days of data.

- *First period* July 5: Germans attack heavily fortified Soviet positions.
- *Second period* July 6-July 11: Germans continue a more organized attack.
- *Third period* July 12: Soviets counterattack when Germans were continuing their attack.
- *Fourth period* July 13-July 18: Soviets attack and Germans make a hasty defense).

A different d parameter is fit to each of the four parts of the campaign using the same a, b, p, q parameters shown in equations IV.B.3.b.(34) and IV.B.3.b.(35) for the data in Table 19. This will be referred to as Model 1 for this section. The results are as follows.

The first period had the smallest SSR value when d=0.91. The second period had the smallest SSR value when d=1.24. The third period is considered to have the tactical parameter d=1 because there was no defender during the third period. The fourth period had the smallest SSR value when d=1.17.

The interpretation of the d values found is that the d value of 0.91 for the first period (i.e. the defender having the advantage), definitely makes sense because the Germans were attacking against the heavily fortified Soviet positions, and as a result, the Soviets inflicted heavier casualties on the Germans than the Germans did on the Soviets.

By intuition, it is likely that Soviets will continue to have the defender advantage through the second period as well. But this is not the case, since d=1.24, meaning that even if it were the Germans attacking they were more advantageous than the Soviets who were in their defensive postures. That is, it was the Soviets who were losing more.

The third period is considered to be the day that neither side is defending, so no interpretation is needed.

The fourth period has a d value of 1.17, which again indicates an attacker advantage. The value 1.17 indicates a slightly smaller attacker advantage than the Germans had during the second period. The Soviets had an attacker advantage during the fourth period, but not one so great as the Germans had during the second period.

The SSR values of the first, second, third and fourth periods mentioned above are 1.93×10^7, 3.70×10^7, 3.83×10^8, 9.53×10^7, respectively. The overall sum of the SSR values is 5.34×10^8 for the whole campaign, which gives a 4% better fit than the previous results. Figures 42 and 43 show the fitted versus real losses for the Soviet and German forces, respectively, for Model 1.

Overall, these results interpreted above indicate that for the Battle of Kursk, other than on the first day, it was always advantageous to be the attacker.

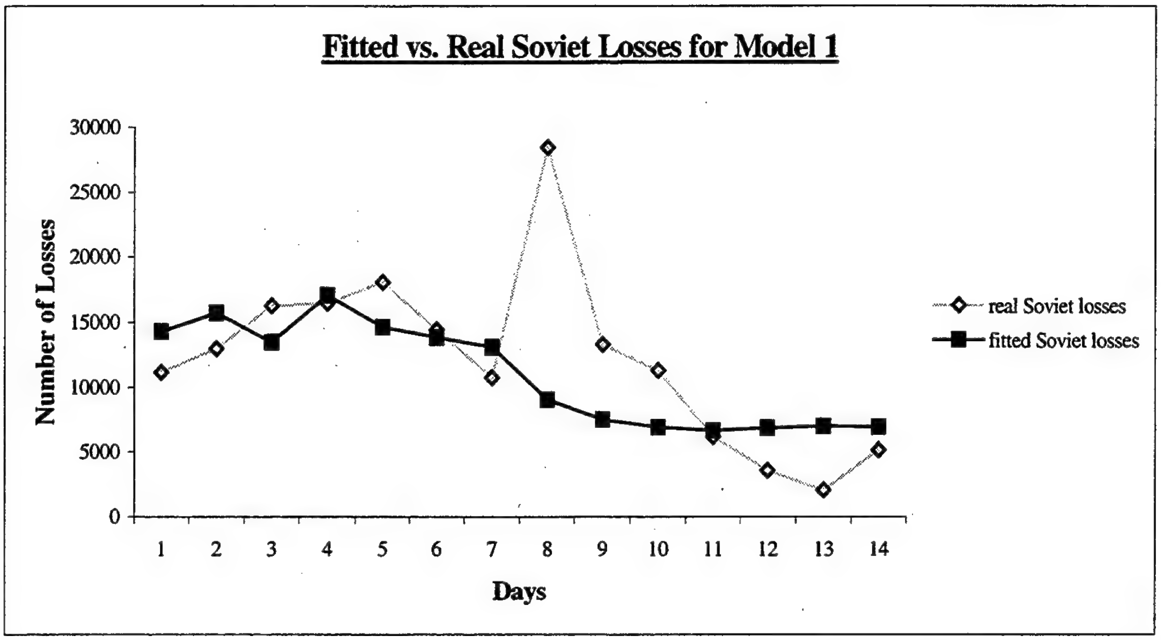

Figure 42. Fitted losses plotted versus real losses for the Soviet forces for model 1, which has four periods, and d=1 for the 8th day of the battle.

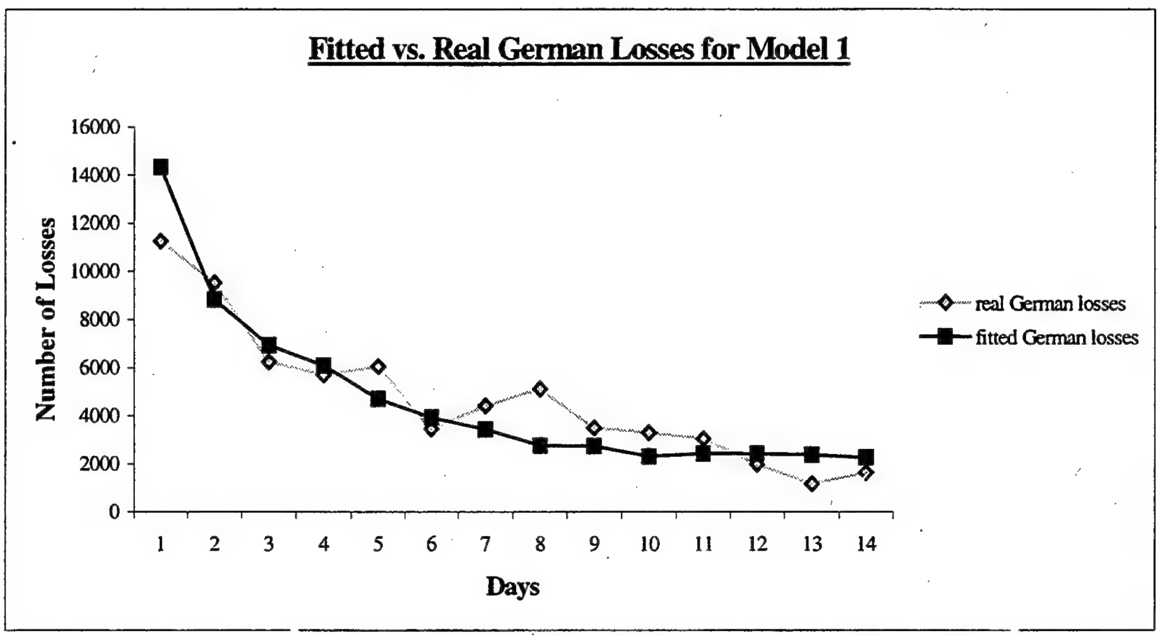

Figure 43. Fitted losses plotted versus real losses for the German forces for model 1, which has four periods, and d=1 for the 8th day of the battle.

One could argue that the third period, having no tactical parameter, does not make sense. If the third period is considered to have a tactical parameter of its own that is independent from the others, assuming that it was the day on which Soviets attacked, it is found to be d=0.32. This result obviously indicates an absolute defender advantage for the Germans and attacker disadvantage for the Russians. This will be referred as Model 2 for this section. In such an approach, the SSR value for the third period will be 1.78×10^7 giving an overall SSR value of 1.69×10^8 — almost a 70% better fit than the result found for Model 1 above. This is a much better fit because the biggest outlier now has its own unique d parameter, and is essentialy removed. This is also a clear indication of the tremendous effect of one outlier on the fit of the models. Figures 44 and 45 show the fitted versus real losses for the Soviet and German forces, respectively for Model 2.

Based on the results above, it can be concluded that considering the campaign in four different parts definitely helps to find a better fit. So, for combat modeling purposes, the tactical parameter values should depend on the situation of the battle.

Another approach is to leave out only the data for July 12, and not to divide the campaign into four periods, (i.e. considering it as a whole, using the same a, b, p, q parameters and fitting a new d parameter under these given circumstances). This model is referred as Model 3 for this section. By following this methodology, d is found to be 1.14 with an SSR value of 1.89×10^8 which is a 12% worse fit than Model 2, but still a 65% better fit than Model 1. In Model 2, a different d parameter for period 3 essentially removed the outlier.

Figures 46 and 47 show the fitted versus real losses for the Soviet and German forces, respectively, for Model 3.

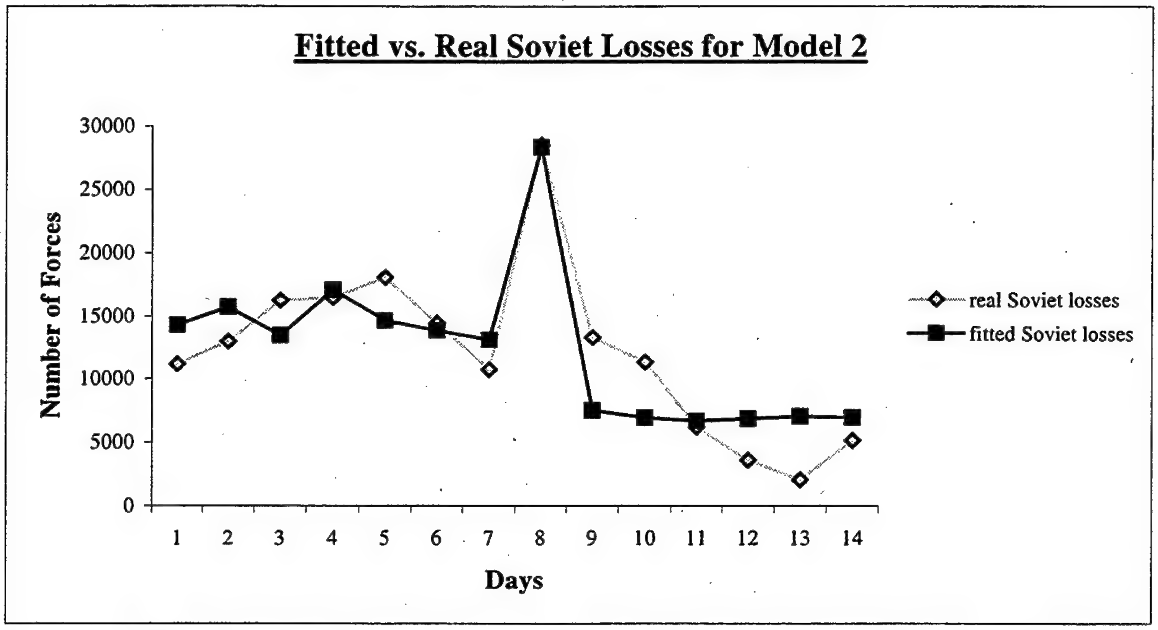

Figure 44. Fitted losses plotted versus real losses for the Soviet forces for model 2, which has four periods, and the Soviets as the attacker for the 8^{th} day of the battle.

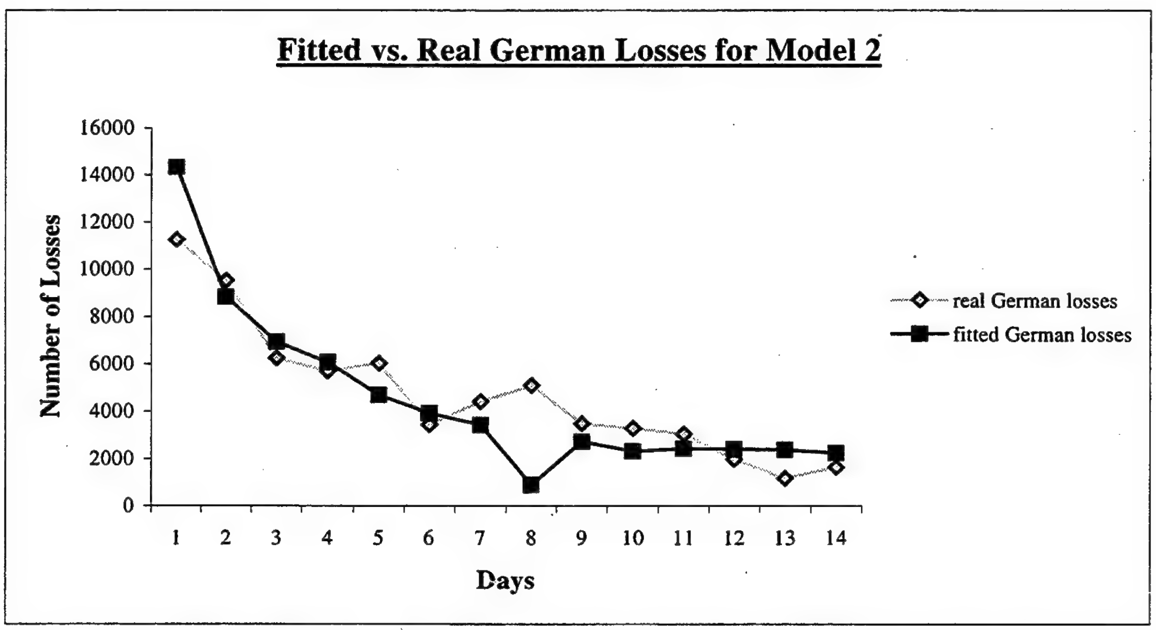

Figure 45. Fitted losses plotted versus real losses for the German forces for model 2, which has four periods, and the Soviets as the attacker for the 8^{th} day of the battle.

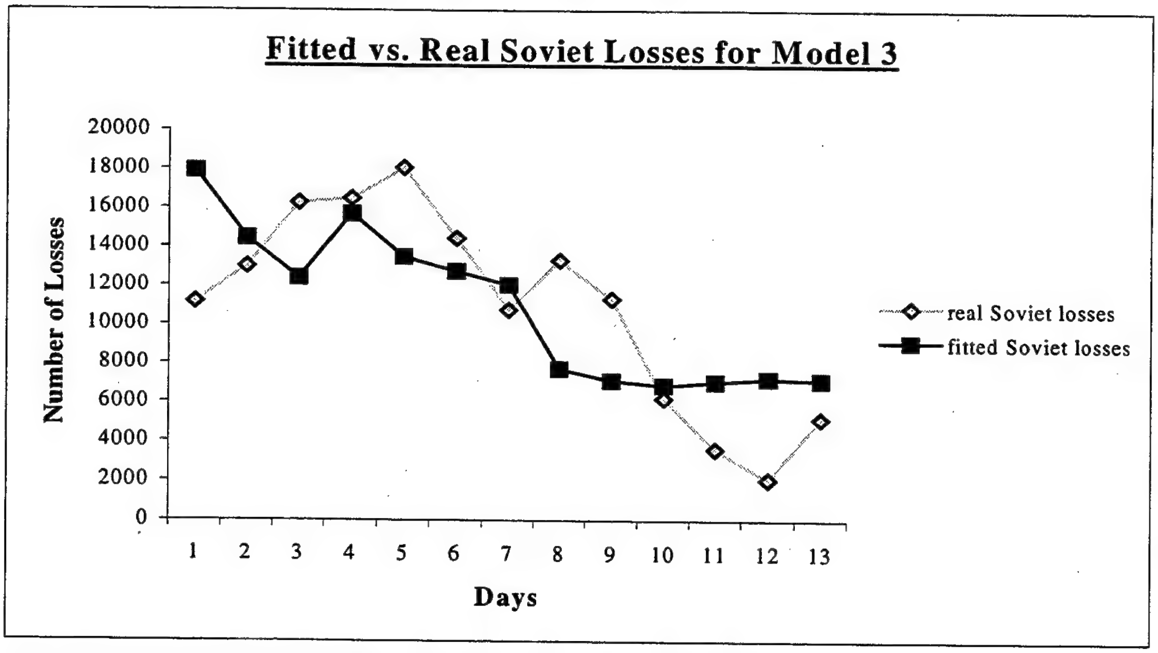

Figure 46. Fitted losses plotted versus real losses for the Soviet forces for Model 3 which leaves out the 8th day of the battle, does not divide the campaign into 4 periods, uses the same parameters as Model 1 and Model 2 and fits a new *d* parameter.

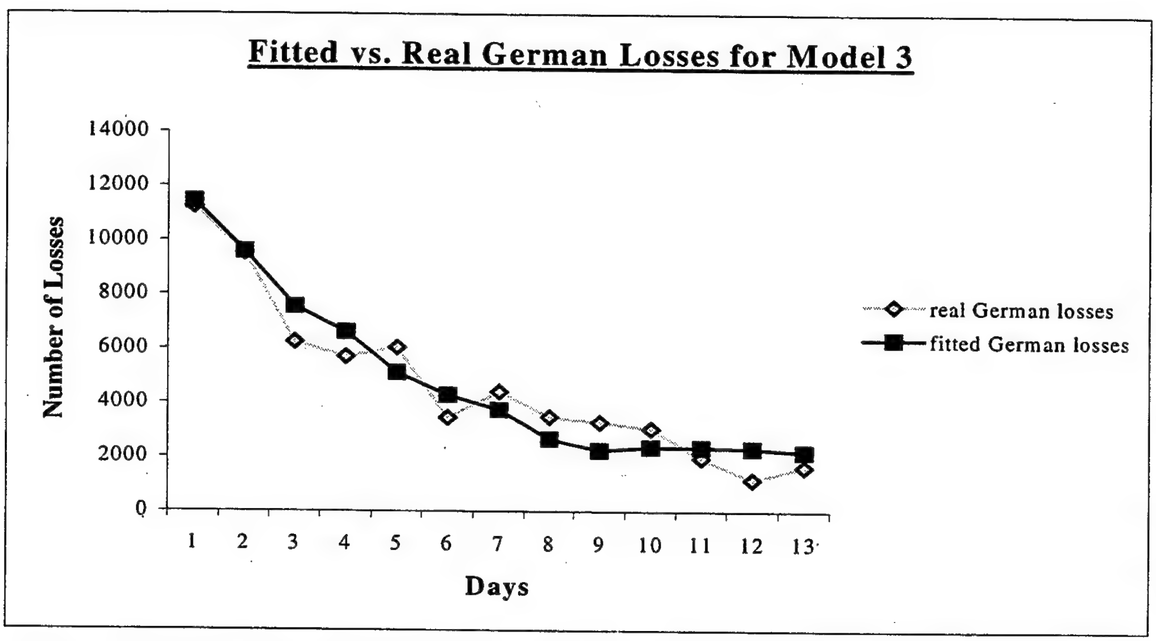

Figure 47. Fitted losses plotted versus real losses for the German forces for Model 3 which leaves out the 8th day of the battle, does not divide the campaign into 4 periods, uses the same parameters as Model 1 and Model 2 and fits a new *d* parameter.

This results in the question; What if a whole new regression analysis is done to the data, leaving out the eighth day? This model is referred to as Model 4 and by doing so, the resulting model with an SSR value of 1.90×10^{8} is found to be:

$$\dot{B} = 1.85\times10^{-51} R^{9.6853} B^{0.1458} \tag{40}$$

$$\dot{R} = 3.56\times10^{-53} B^{9.6853} R^{0.1458} \quad . \tag{41}$$

These results are far better than those found in previous sections that contained the outlier. But, they do not however, provide a better fit than the ones found in this section which are adjusted for the outlier. Also, it is significant that there is a big difference in the size of the p and q parameters. Figures 48 and 49 show the fitted versus real losses for the Soviet and German forces, respectively, for model 4.

Handling the data in parts and fitting different tactical parameters definitely improves the fits of all models given in this section. This result is consistent with what Hartley and Helmbold found in their studies [Ref.10].

Model 2 with an SSR value of 1.69×10^{8} has the smallest SSR value thus far. This result largely depends on considering July 12, which is the largest outlier apart from the rest of the data, causing a considerable decrease from the previous lowest SSR value of 5.54×10^{8} to a much lower SSR value of 1.69×10^{8}.

Model 3 finds d to be 1.14, which means an attacker advantage/defender disadvantage. But, this circumstance again largely depends on still using the same parameters that we had when the tactical parameter d is 1.17. Once more, this d value indicates an attacker advantage/defender disadvantage situation.

In Model 4, leaving even only one day out (the largest outlier), improves the model's fit tremendously when compared to the previous SSR values.

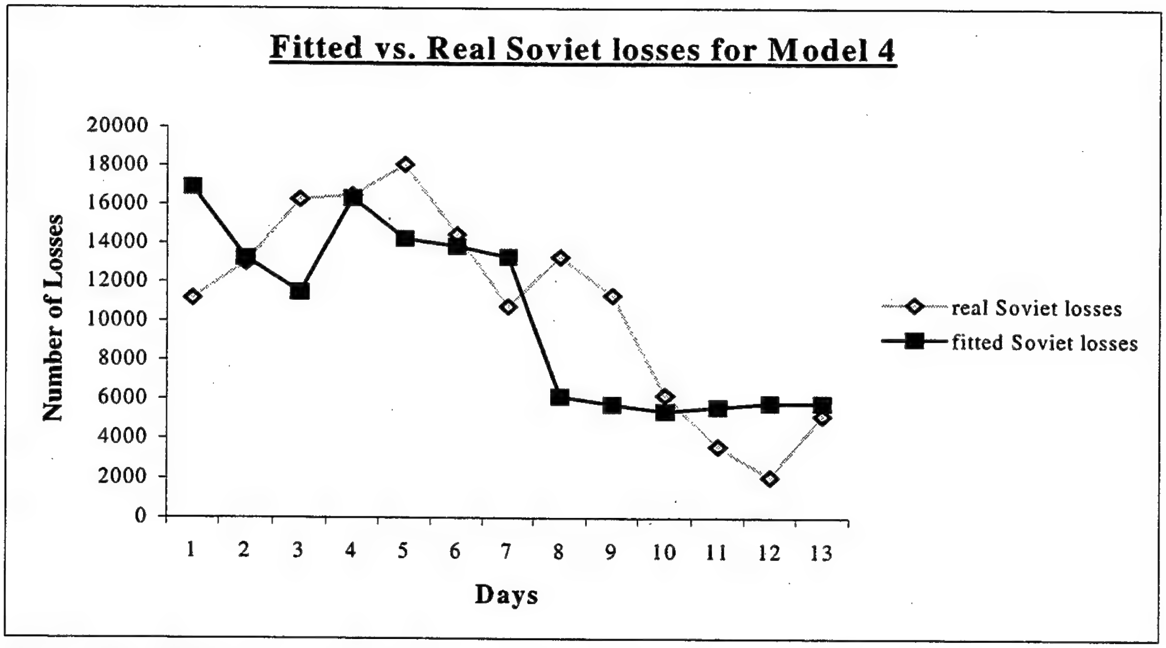

Figure 48. Fitted losses plotted versus real losses for the Soviet forces for model 4, which leaves out the 8th day of the battle, does not divide the campaign into 4 periods, fits a whole new regression model.

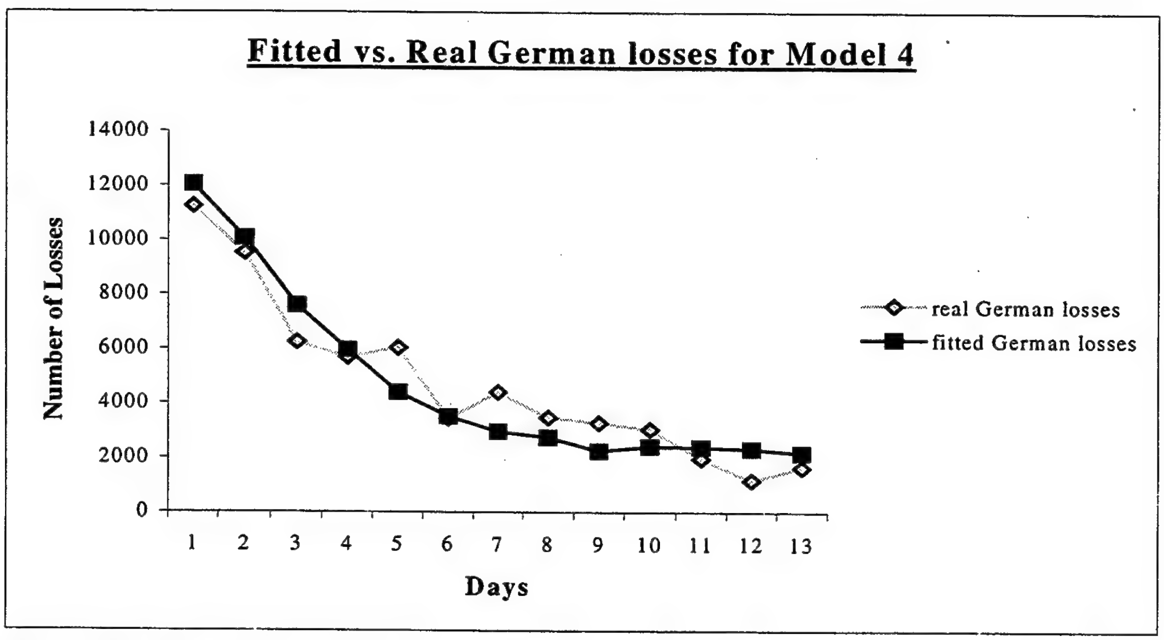

Figure 49. Fitted losses plotted versus real losses for the German forces for model 4, which leaves out the 8th day of the battle, does not divide the campaign into 4 periods, fits a whole new regression model.

For all models, a and b parameters are significantly small and $a>b$. This result suggests that individual German effectiveness was greater than individual Russian effectiveness.

When the p and q parameters are compared, it is observed that the p parameter is greater than the q parameter suggesting that one side's losses are more a function of his own forces rather than being a function of the opponent's forces. This observation is different from what Fricker found in his study.

The results for the four different models are given in Table 30.

Name of the model	a	b	p	q	d	SSR	R^2
Campaign in four Parts	1.88E-47	1.07E-48	7.5038	1.5793	4 periods d=0.91,1.24, 1.0,1.17	5.34E+8	-2.3410
Campaign in four Parts	1.88E-47	1.07E-48	7.5068	1.5793	4 periods d=0.91,1.24, 0.32,1.17	1.69E+8	-0.0607
Campaign in four Parts	1.88E-47	1.07E-48	7.5038	1.5793	1.14	1.89E+8	0.5689
Campaign in four Parts	1.85E-51	3.56E-53	9.6853	0.1458	-	1.90E+8	0.5658

Table 30. The results for the model which considers the battle in separate parts.

The negative R^2 values are mainly a result of considering certain days in the campaign solely on their own. This results in SST value for that day being zero. This result (i.e, the SST value being zero for a certain day) is the main reason for negative R^2 values in this section.

5. Considering change points in the model

The findings in the previous sections suggest that fitting models for separate phases of the battle might improve the fit to the data. This section considers one or more

attrition change points for each side. At each chosen point in the phase of battle all the parameters pertaining to that particular side will change.

When the historical account of the battle is taken into account, it is apparent that the Germans generally attacked between July 5, and July11, for the first seven days, and the Soviets attacked between the days July 12, and July 18, for the last seven days. This is the first change point to be considered and will be referred as change point 7/7.

Another approach is considering that the Germans attacked between July 5, and July 12, for the first eight days, and the Soviets attacked between July 13, and July 18, for the last six days. This is the second change point to be considered, and will be referred to as change point 8/6. This type of approach (considering change points for fitting the model to the data) is similar to what Hartley and Helmbold did in their study [Ref.10].

No tactical parameter will be considered, and only linear regression will be used in fitting the data to the model with change points. For estimating the parameters of the model that minimize the sum of squared residuals of the actual and estimated attrition, S-PLUS software and the last 14 days of the aggregated data given in Table 14 in Section IV.A.1 are used.

Results for the first half of the Linear Regression model for change point 7/7 with an SSR value of 6.53×10^{7} are:

$$\dot{B} = 8.91\times10^{-30} R^{6.4117} B^{-0.4323} \tag{42}$$

$$\dot{R} = 2.62\times10^{-31} B^{6.4117} R^{-0.4323} \tag{43}$$

Results for the second half of the Linear Regression model for change point 7/7 with an SSR value of 8.78×10^{7} are:

$$\dot{B} = 1.90 \times 10^{-292} R^{18.0587} B^{34.4502} \tag{44}$$

$$\dot{R} = 4.37 \times 10^{-291} B^{18.0587} R^{34.4502} \tag{45}$$

where both halves add up to a total SSR value of 1.53×10^8, and result in an R^2 value of 0.7448.

Results for the first half of the Linear Regression model for change point 8/6 with an SSR value of 1.65×10^8 are:

$$\dot{B} = 7.75 \times 10^{-5} R^{4.4212} B^{-2.8454} \tag{46}$$

$$\dot{R} = 1.91 \times 10^{-6} B^{4.4212} R^{-2.8454} \tag{47}$$

Results for the second half of the Linear Regression model for change point 8/6 with an SSR value of 7.78×10^7 are:

$$\dot{B} = 1.94 \times 10^{-246} R^{25.7652} B^{18.7674} \tag{48}$$

$$\dot{R} = 1.32 \times 10^{-247} B^{25.7652} R^{18.7674} \tag{49}$$

where both halves add up to a total SSR value of 2.43×10^8 and result in an R^2 value of 0.3488.

The SSR value for the change point 7/7 is the smallest SSR value we have seen. It gives a 9% better fit than Model 2 of Section IV.B.4 which is 1.69×10^8. It is also almost a 56% better fit than the one found in section IV.B.1, where only one set of parameters is fit to the whole data. This model has the highest R^2 value we have seen thus far, and easily the best fit we have obtained. We can conclude that fitting the model using the change points definitely improves the fit, and this is consistent with the result Hartley and Helmbold [Ref. 10] found in their study.

However the only concern is that the q parameter for both the change point 7/7 and change point 8/6 are negative, meaning that the number of a force's casualties

decreases as one of the force strengths increases. The p and q parameters found in the models are extremely high. Doubling the force size results in a dramatic change in the outcome and this does not intuitively make sense. Since this analogy is both illogical and unlikely, we resolve that even if the change point approach gives the lowest SSR value of 1.53×10^{8}, with the change point 7/7 model, we cannot accept this fit as the best one. This result also suggests a wide range of parameters gives similar fits to the data

In all the models explored in this section, the a and b parameters are significantly small, and except the equations given in IV.B.5.(44), IV.B.5.(45), $a>b$. This suggests that individual German effectiveness was greater than individual Russian effectiveness.

When the p and q parameters are compared, it is observed that except the model given in equations IV.B.5.(44), IV.B.5.(45), the p parameter is greater than the q parameter. This comparison suggests that one side's losses are more a function of his own forces rather than being a function of the opponent's forces, and is different from what Fricker found in his study.

6. Using different weights

This section considers different weights for aggregating the battle data. Bracken [Ref.8] states in his study that, "The given weights are consistent with those of studies and models of the U.S.Army Concepts Analysis Agency. Virtually all theater-level dynamic combat simulation models incorporate similar weights, either as inputs or as decision parameters computed as the simulations progress." Although Bracken's points are well taken, this study will try to fit models by using different weights for exploratory purposes. The different weights are selected on a wholly intuitive basis and are a result of many different trial and error calculations.

The first weight combination will use the weights 1, 5, 20 and 40; the second weight combination will use the weights 1, 5, 15 and 20; the third weight combination will use the weights 1, 5, 30 and 40; the fourth weight combination will use the weights 1, 5, 20 and 30 for manpower, APC, artillery and tanks, respectively.

Note that tanks are weighted more because the Battle of Kursk was a major tank battle. Both linear and robust LTS regression models are used to fit the data, which is aggregated using the different weight combinations given above.

Table 31 presents the aggregated data obtained using the first weight combination. Table 32 presents the aggregated data obtained using the second weight combination. Table 33 presents the aggregated data obtained using the third weight combination. Table 34 presents the aggregated data obtained using the fourth weight combination.

4. *First weight combination*

The result for the linear regression model that gives an SSR value of 1.15×10^{9} and an R^2 value of 0.0870, is:

$$\dot{B} = 1.25\times10^{-38} R^{5.2298} B^{2.2746} \quad (50)$$

$$\dot{R} = 1.60\times10^{-39} B^{5.2298} R^{2.2746} \quad (51)$$

The result for the robust LTS regression model that gives an SSR value of 1.07×10^{9} and an R^2 value of 0.1514, is:

$$\dot{B} = 7.26\times10^{-35} R^{5.5312} B^{1.3268} \quad (52)$$

$$\dot{R} = 5.53\times10^{-36} B^{5.5312} R^{1.3268} \quad (53)$$

b. *Second weight combination*

The result for the linear regression model that gives an SSR value of 6.24×10^{8} and an R^2 value of 0.0975, is:

Day	Blue Forces	Blue Losses	Red Forces	Red Losses
1	620173	13007	369811	14737
2	609589	14733	356025	14392
3	587405	21146	349465	8529
4	567452	22492	360184	7602
5	545652	23671	353044	8703
6	531943	17325	353484	3930
7	522036	13314	352210	5375
8	487313	36452	348085	6832
9	476711	15442	348389	4491
10	465684	13583	345740	4110
11	468799	8021	342382	4567
12	470151	4080	345266	2355
13	468706	2907	347745	1274
14	465083	6780	348400	1539

Table 31. Data on forces which are aggregated by using weight combination 1. Forces are combat manpower, APCs, artillery and tanks, which are weighted by 1, 5, 20 and 40 respectively. Here, a tank is considered to be twice as valuable as an artillery piece.

Day	Blue Forces	Blue Losses	Red Forces	Red Losses
1	568728	10842	344261	10657
2	558869	12243	335240	9407
3	542820	15891	330235	6074
4	529132	16122	342199	5377
5	512552	17846	337194	5893
6	500678	14120	336529	3150
7	491876	10579	334920	4040
8	464708	28092	332115	4812
9	454121	13052	332554	3316
10	444529	11198	330220	3165
11	448134	6076	327947	2972
12	447451	3525	329471	1875
13	446136	2067	330695	1124
14	443168	5060	330730	1364

Table 32. Data on forces which are aggregated by using weight combination 2. Forces are combat manpower, APCs, artillery and tanks, which are weighted by 1, 5, 15 and 20 respectively. Here a tank is considered to be 33% more valuable than artillery.

Day	Blue Forces	Blue Losses	Red Forces	Red Losses
1	627223	13137	381471	14977
2	616349	15033	367635	14442
3	594015	21296	361005	8599
4	573932	22632	372314	7732
5	552052	23761	365144	8763
6	538233	17455	365474	4050
7	528316	13384	364270	5525
8	493443	36612	360025	6952
9	482771	15542	360259	4561
10	471714	13633	357580	4160
11	474809	8071	354212	4597
12	476151	4110	357056	2395
13	474726	2907	359565	1294
14	470993	6820	360220	1649

Table 33. Data on forces which are aggregated by using weight combination 3. Forces are combat manpower, APCs, artillery and tanks, which are weighted by 1, 5, 30 and 40 respectively. Here an artillery piece is considered to be six times more effective than an APC and a tank is considered to be eight times more valuable than an APC.

Day	Blue Forces	Blue Losses	Red Forces	Red Losses
1	596213	11957	359951	12757
2	585919	13563	348535	11912
3	566765	18556	342735	7319
4	549912	19342	354224	6522
5	530702	20781	348144	7313
6	517883	15755	348004	3570
7	508526	11964	346580	4745
8	477543	32312	343085	5852
9	466931	14272	343439	3921
10	456614	12403	340940	3650
11	459969	7061	338122	3777
12	460301	3810	340316	2125
13	458926	2487	342175	1204
14	455603	5930	342520	1479

Table 34. Data on forces which are aggregated by using weight combination 4. Forces are combat manpower, APCs, artillery and tanks, which are weighted by 1, 5, 20 and 30 respectively. Here an artillery piece is considered to be four times more effective than an APC and a tank is considered to be six times more valuable than an APC.

$$\dot{B} = 2.50\times10^{-46} R^{5.7638} B^{3.1222} \tag{54}$$

$$\dot{R} = 3.49\times10^{-47} B^{5.7638} R^{3.1222} \tag{55}$$

The result for the robust LTS regression model that gives an SSR value of 5.48×10^{8} and an R^2 value of 0.2072, is:

$$\dot{B} = 7.85\times10^{-36} R^{5.8613} B^{1.1899} \tag{56}$$

$$\dot{R} = 4.75\times10^{-37} B^{5.8613} R^{1.1899} \tag{57}$$

c. Third weight combination

The result for the linear regression model that gives an SSR value of 1.15×10^{9} and an R^2 value of 0.0926, is:

$$\dot{B} = 3.78\times10^{-39} R^{5.2293} B^{2.3513} \tag{58}$$

$$\dot{R} = 5.34\times10^{-40} B^{5.2293} R^{2.3513} \tag{59}$$

The result for the robust LTS regression model that gives an SSR value of 1.06×10^{9} and an R^2 value of 0.1637, is:

$$\dot{B} = 1.46\times10^{-35} R^{5.9619} B^{1.0159} \tag{60}$$

$$\dot{R} = 9.33\times10^{-37} B^{5.9619} R^{1.0159} \tag{61}$$

d. Fourth weight combination

The result for the linear regression model that gives an SSR value of 8.63×10^{9} and an R^2 value of 0.0943, is:

$$\dot{B} = 2.89\times10^{-42} R^{5.4863} B^{2.666} \tag{62}$$

$$\dot{R} = 3.91\times10^{-43} B^{5.4863} R^{2.666} \tag{63}$$

The result for the robust LTS regression model that gives an SSR value of 7.74×10^{8} and an R^2 value of 0.1873, is:

$$\dot{B} = 5.05\times10^{-35} R^{5.6294} B^{1.2631} \tag{64}$$

$$\dot{R} = 3.51\times10^{-36} B^{5.6294} R^{1.2631} \tag{65}$$

Using different weights to aggregate the data can improve the fit to the data. The SSR value observed for the second weight combination when the data is fitted using the Robust LTS Regression model is the lowest SSR value found for models without the tactical parameter *d*. But, this result may be due to the small size of the weights used for aggregating the data. Comparing SSR values makes sense as long as the weights used for aggregating the data are constant for all models compared, but this is not the case in our discussion. In such circumstances, the R^2 value is a better parameter to use for comparison purposes rather than the SSR value because the R^2 value adjusts to scale. How the R^2 value is computed is given in equation IV.A.1.b.(10).

The parameters and the R^2 values for each weight combination are given in Table 34 for both linear regression and robust LTS regression models. When the R^2 values are compared for the models presented in this section, it is observed that weight combination 2 gives the best fit when the robust LTS regression technique is used, with the greatest R^2 value of 0.2072. The second best fit is found when weight combination 4 is used with the robust LTS regression technique, and the third best fit is found when weight combination 3 is used, again with robust LTS regression technique. These models with different weight combinations do not give a better fit as a whole when compared to the two models given in IV.B.1.d.(26), IV.B.1.d.(27) and IV.B.3.d.(38), IV.B.3.d.(39) where

both models have an R^2 value of 0.2262 and use the weight combination of 1, 5, 20 and 40, for combat manpower, APCs, tanks and artillery, respectively.

When the p and q parameters are compared, it is evident that for all the models discussed in this section, the p parameter is greater than the q parameter. This result suggests that one side's losses are more a function of the opponent's forces rather than being a function of his own forces, resembling earlier findings.

Except for the model given in IV.B.6.d.(64) and IV.B.6.d.(65), the a and b parameters are significantly small and $a>b$ for all the models discussed in this section. This result suggests that individual German effectiveness was greater than individual Russian effectiveness.

One can easily argue that tanks are more effective during an offensive then they are during a defense. Likewise, artillery can be considered to have different effects on the outcome of the battle depending on the type of a campaign. The weights used in the second weight combination may give a better fit than the models which use the other three weight combinations. However the relevance of the weights used is another topic of discussion in itself. In short, it is clear according to our examples that changing the weights can help find a better fit, but one must be careful in doing so that the issue of relevancy to the real world is not ignored. Further investigation is recommended for determining weight combinations.

Figures 49 and 50 show fitted losses plotted versus real losses for the Soviet and the German forces respectively, for the robust LTS regression model using the second weight combination which gives the best fit.

For ease of comparison, the results for all the models using different weight combinations and the previous two results are given in Table 35.

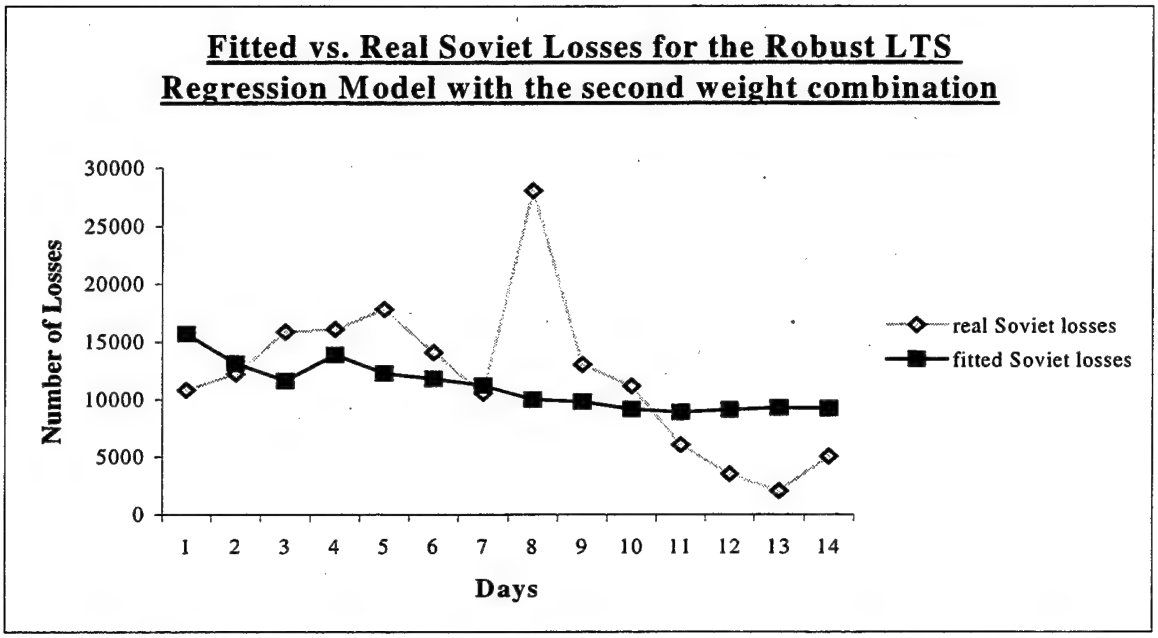

Figure 50. Fitted losses plotted versus real losses for the Soviet Forces for the robust LTS regression model using the weight combination 2. The same pattern where the model over/underestimates the battle in three distinctive phases is observable in this plot.

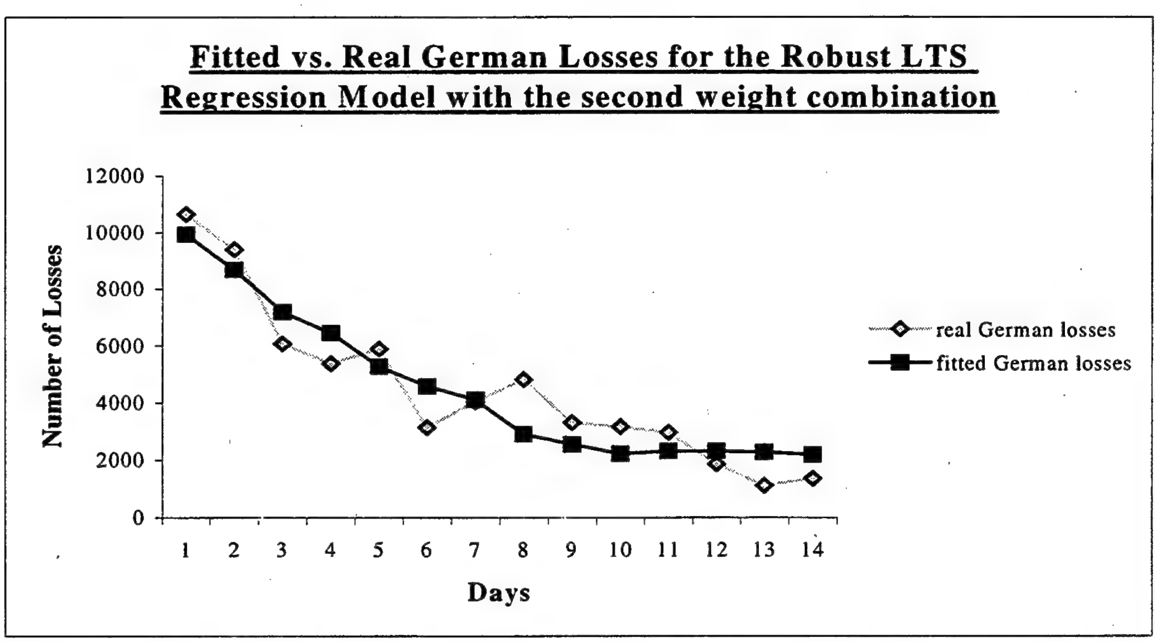

Figure 51. Fitted losses plotted versus real losses for the German Forces for the robust LTS regression model with the weight combination 2.

Type Of the model	*a*	*b*	*p*	*q*	*d*	*SSR*	R^2
Previous Best Result	2.27E-40	1.84E-41	6.0843	1.7312	-	5.54E+8	0.2262
Weight Comb.1 Lin.Reg.	1.25E-38	1.60E-39	5.2298	2.2746	-	1.15E+9	0.0870
Weight Comb.1 Rob.LTS	7.26E-35	5.53E-36	5.5312	1.3268	-	1.07E+9	0.1514
Weight Comb.2 Lin.Reg.	2.50E-46	3.49E-47	5.7638	3.1222	-	6.24E+8	0.0975
Weight Comb.2 Rob.LTS	7.85E-36	4.75E-37	5.8613	1.1899	-	5.48E+8	0.2072
Weight Comb.3 Lin.Reg.	3.78E-39	5.34E-40	5.2293	2.3513	-	1.15E+9	0.0926
Weight Comb.3 Rob.LTS	1.46E-35	9.33E-37	5.9619	1.0159	-	1.06E+9	0.1637
Weight comb.4 Lin.Reg.	2.89E-42	3.91E-43	5.4863	2.6660	-	8.63E+9	0.0943
Weight Comb.4 Rob.LTS	5.05E-35	3.51E-36	5.6294	1.2631	-	7.74E+8	0.1873

Table 35. The results for the models using different weight combinations. Weight combination 2 gives the best fit.

7. Force ratio and fractional exchange ratio models

In this section, Force Ratio (FR) and Fractional Exchange Ratio (FER) models are explored and analyzed. The reason for including this approach in our discussion is that both analysts and military staff use force ratios in models for combat outcomes and decisions. For this purpose, five different models are investigated. The first model uses the FR of aggregated forces as a predictor to predict the percent of casualties for each side. The FR of blue forces is equal to the total number of aggregated blue forces divided by the total number of aggregated red forces, and likewise for the FR of the red forces.

The percent of casualties of the blue forces is equal to the total number of aggregated blue losses divided by the total number of aggregated blue forces.

Figures 52 and 53 show loss ratio plotted against the FR for Soviet and German forces, respectively. The representation of Model 1 looks like:

$$(\dot{B}/B) = I_1 + I_2 + (B/R) \tag{66}$$

$$(\dot{R}/R) = I_1 + I_2 + (R/B) \tag{67}$$

where I_1 is an indicator of the blue force or red force, and I_2 indicates the difference between the attacker and defender, and are given as:

$$\begin{aligned} I_1 &= 1 \text{ if Blue} \\ I_1 &= 0 \text{ if Red} \end{aligned} \tag{68}$$

$$\begin{aligned} I_2 &= 1 \text{ if attacker} \\ I_2 &= 0 \text{ if defender} \end{aligned} \tag{69}$$

The resulting model for Model 1 with the intercept that gives an SSR value of 3.09×10^{-3} and an R-squared value of 0.2296 (given by the S-PLUS software) is:

$$PC = -0.0103 - 0.0074 I_1 + 0.0068 I_2 + 0.0275(OFR) \tag{70}$$

where PC denotes the percent of casualties as given in IV.B.7.(64), IV.B.7.(65), and OFR denotes the opponent's FR for a given side.

The R-squared value given above is not calculated using the formula given in equation IV.A.1.(10) but given by the S-PLUS software and will be used for all the models throughout this section.

Here, indicator variables are mainly used for the purpose of adjusting the intercept. When the intercept term is used in the model, the correlation matrix of the estimated coefficients shows a high correlation between the estimates that, due to high

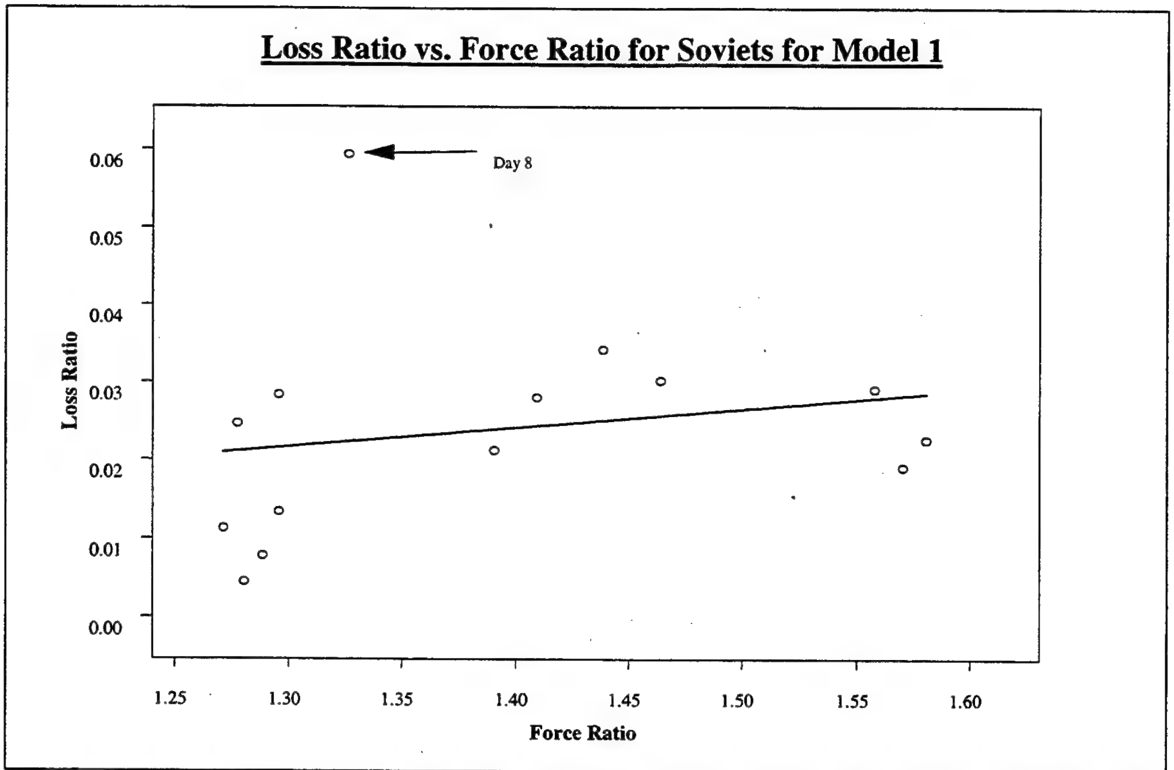

Figure 52. Loss ratio plotted versus force ratio for Soviet forces for model 1. Soviets lost a higher percentage of their forces as their force ratio increased.

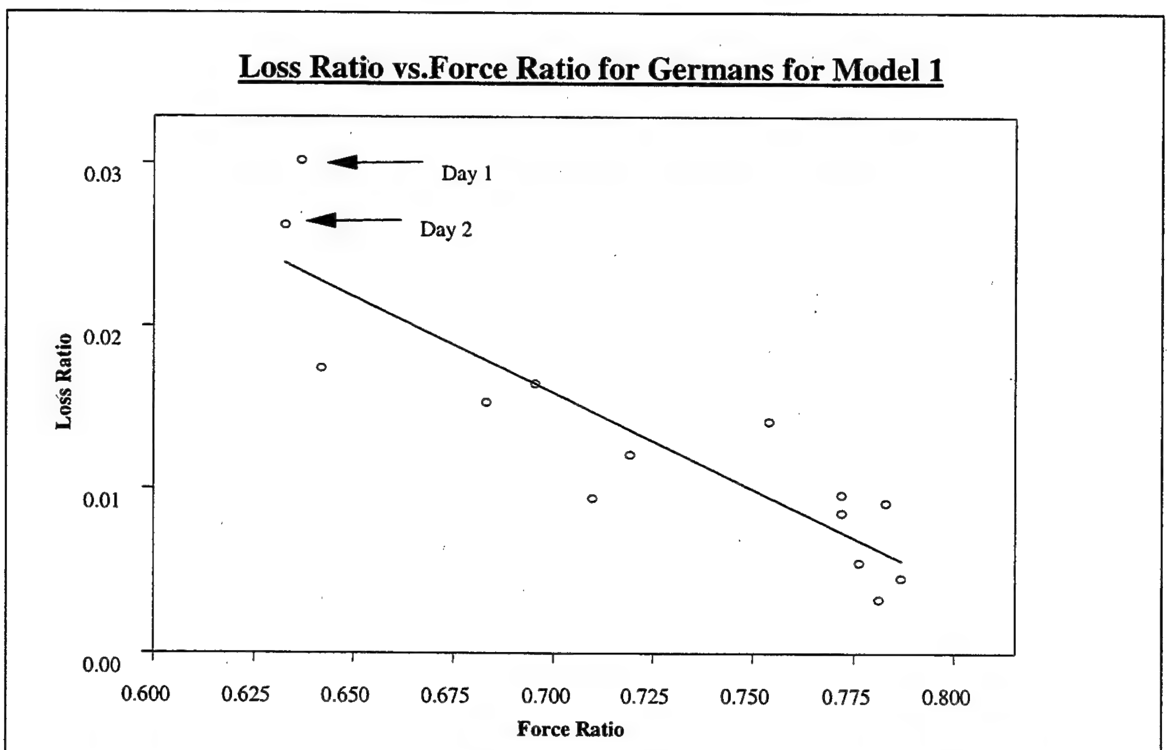

Figure 53. Loss ratio plotted versus force ratio for German forces for model 1. Germans lost a lower percentage of their forces as their force ratio increased.

	Intercept	I_1	I_2
I_1	0.9754		
I_2	-0.8662	-0.8291	
	-0.9945	-0.9895	0.8379

Table 36. Correlation matrix of the estimated coefficients of the first model. Notice the high correlation between the model's coefficients, and especially the correlation between the intercept and the force ratio, which can result in a very bad fit.

collinearity can result in very inaccurate estimates [Ref.18]. Because of this result, an intercept term is not used in the following models. The correlation matrix of the estimated coefficients is given in Table 36.

Concern over whether or not leaving the intercept term out is correct or not can be addressed by doing a hypothesis test. The null hypothesis will be, H_0: *intercept = 0*, and the alternative hypothesis will be, H_a: *intercept≠0*. With a significance level of $\alpha = 0.1$ and 24 degrees of freedom, the null hypothesis will be rejected if $t \geq t_{0.05,24} = 1.711$ or if $t \leq -t_{0.05,24} = -1.711$. The t-statistics of the intercept of Model 1 is $t = -0.2899$ which is not in the rejection region. So, the null hypothesis is not rejected, and the intercept will be assumed to be zero throughout the models.

The resulting model for Model 1 without the intercept gives an SSR value of 3.105×10^{-3} and a multiple R-squared value of 0.7699 and looks like:

$$PC = 0.001 I_1 + 0.004\varepsilon I_2 + 0.0147(OFR) \tag{71}$$

Table 37 shows the coefficients, standard errors, and t values for Model 1.

	Value	Std. Error	t value	Pr(>|t|)
I1	0.0010	0.0064	0.1534	0.8793
I2	0.0048	0.0039	1.2385	0.2270
OFR	0.0147	0.0045	3.2419	0.0034

Table 37. Important statistical values of the estimated coefficients for Model 1.

The positive coefficient of the indicator variable I_1 indicates a German advantage (though insignificant), where the positive coefficient of the indicator variable I_2 indicates a defender advantage, and again is insignificant. The positive coefficient of the force ratio variable indicates that as the force ratio increases, so do the losses. Even though statistically significant, this result does not intuitively make much sense.

The second model uses the total aggregated force ratios as a predictor to predict the fractional exchange ratios for each side. FER for the blue forces is equal to the percent of blue casualties divided by the percent of red casualties, and likewise for the FER of the red forces. Figures 53 and 54 show the FER plotted against force ratio for Soviet and German forces, respectively. The representation of Model 2 looks like:

$$(\dot{B}/B)/(\dot{R}/R) = I_1 + I_2 + (B/R) \tag{72}$$

$$(\dot{R}/R)/(\dot{B}/B) = I_1 + I_2 + (R/B) \tag{73}$$

where I_1 indicates the difference between the blue force and red force, and I_2 indicates the difference between attacker and defender (time of battle) and have the values given in IV.B.7.(68) and IV.B.7.(69).

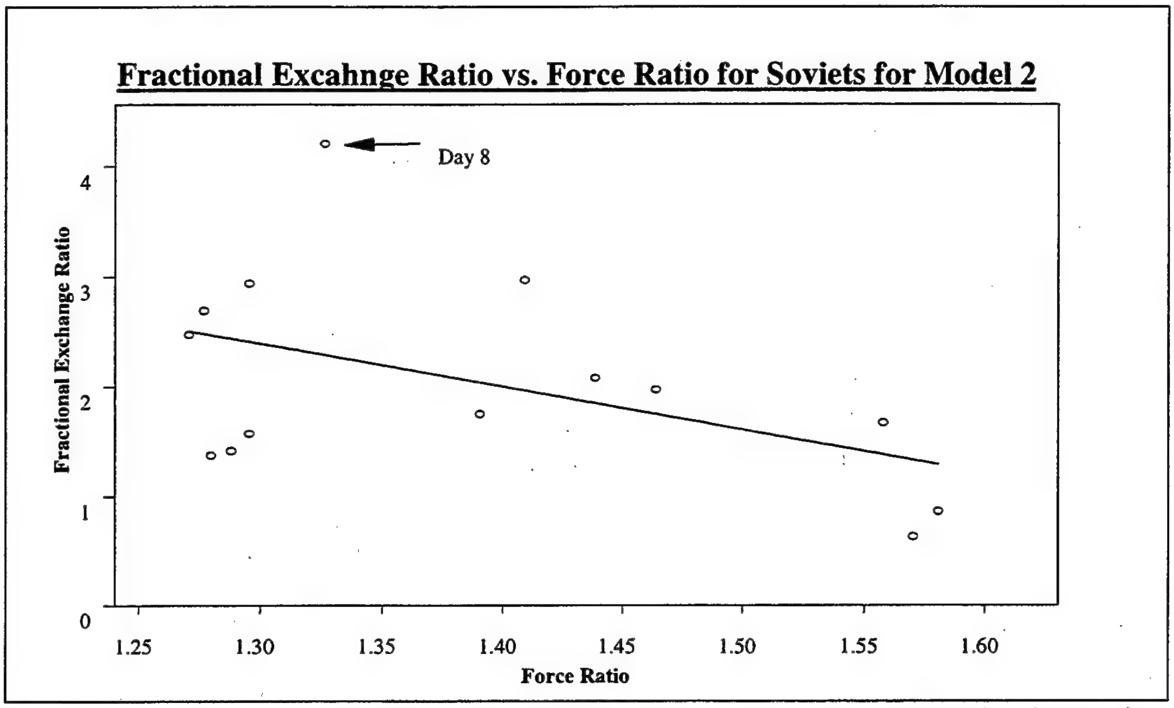

Figure 54. Fractional exchange ratio plotted versus force ratio for Soviet forces for model 2.

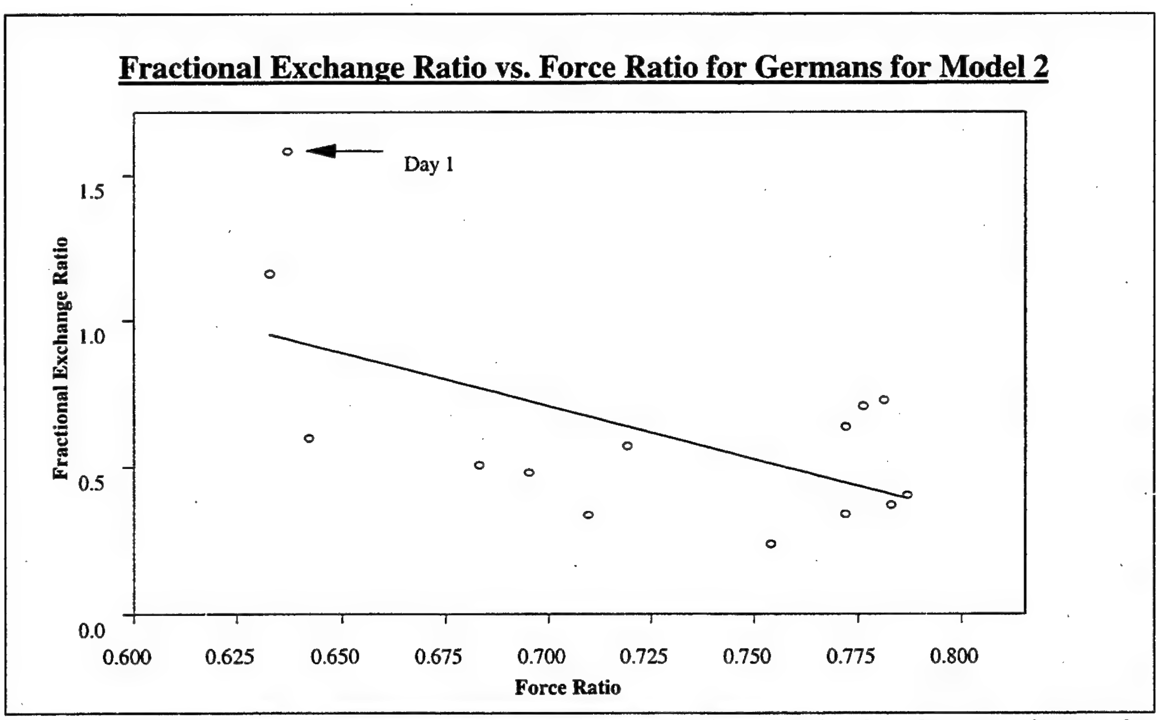

Figure 55. Fractional exchange ratio plotted versus force ratio for German forces for model 2.

Table 38 shows the coefficients, standard errors, and t values for Model 2.

	Value	**Std. Error**	**t value**	**Pr(>|t|)**
I1	1.1849	0.4019	2.9483	0.0068
I2	0.5647	0.2409	2.3441	0.0273
OFR	0.4153	0.2837	1.4638	0.1557

Table 38. Important statistical values of the estimated coefficients for model 2.

The resulting model for Model 2 that gives an SSR value of 12.120 and a multiple R-squared of 0.6963, is:

$$FER = 1.1849 I_1 + 0.5647 I_2 + 0.4153(OFR) \quad (74)$$

where FER denotes the fractional exchange ratio as given in IV.B.7.(72), IV.B.7.(73), and OFR denotes the opponent's FR for a given side.

Similar to the results found for Model 1, the positive coefficient of indicator variable I_1 indicates a German advantage and is significant, where the positive coefficient of indicator variable I_2 indicates a defender advantage and is significant too. The positive coefficient of the force ratio variable indicates that as the force ratio increases so do the losses. Again, the coefficient is not significant and does not intuitively make much sense.

Model 3 uses the force ratio of tanks as a predictor to predict the percent of tank losses for each side. Figures 56 and 57 show the tank loss ratio plotted against the tank force ratio for Soviet and German forces, respectively. The representation of Model 3 looks like:

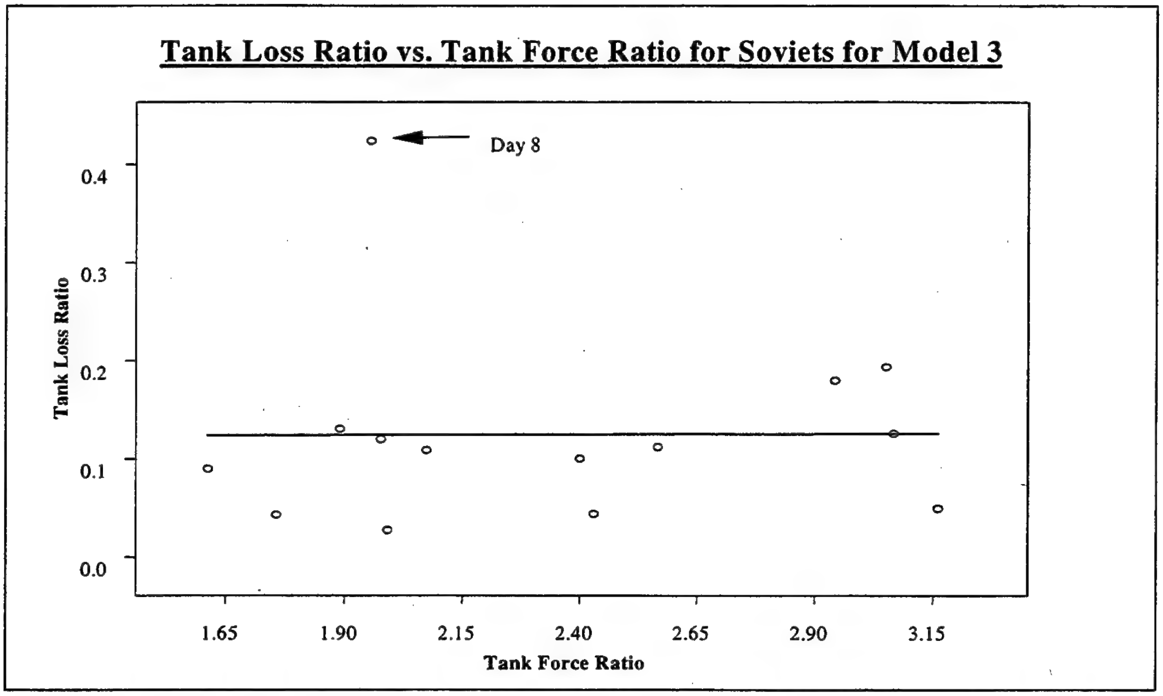

Figure 56. Tank loss ratio plotted versus tank force ratio for Soviet forces for model 3.

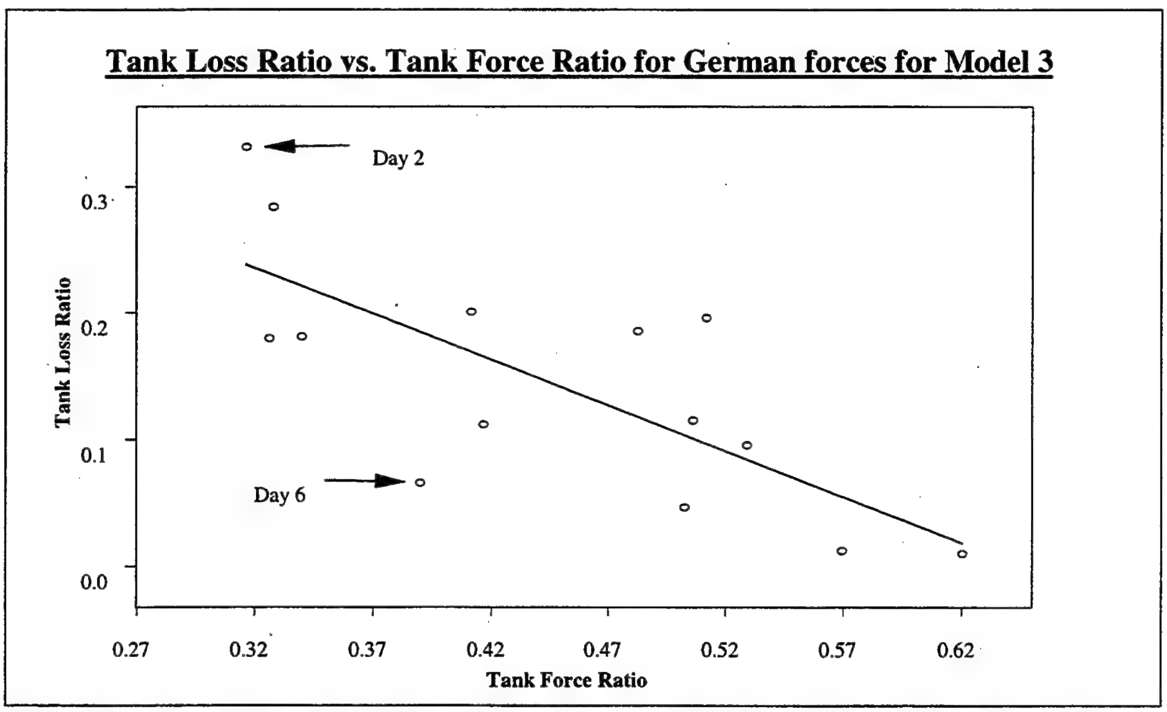

Figure 57. Tank loss ratio plotted versus tank force ratio for German forces for model 3.

$$(BTL / BT) = I_1 + I_2 + (BT / RT) \tag{75}$$

$$(RTL / RT) = I_1 + I_2 + (RT / BT) \tag{76}$$

where I_1 indicates the difference between blue force and red force, and I_2 indicates the difference between attacker and defender (time of battle) and have the values given in IV.B.7.(68) and IV.B.7.(69). BTL, RTL, BT and RT denote blue tank loss, red tank loss, number of blue tanks and number of red tanks, respectively.

Table 39 shows the coefficients, standard errors and t values for Model 3.

	Value	**Std. Error**	**t value**	**Pr(>|t|)**
I1	-0.2703	0.092	-2.9377	0.007
I2	0.1442	0.0291	4.9549	0
OTFR	0.1375	0.0367	3.747	0.0009

Table 39. Critical statistical values of the estimated coefficients of model 3.

The resulting model for model 3, which gives an SSR value of 0.220 and a Multiple R-Squared value of 0.7077 is:

$$PTL = -0.2703 I_1 + 0.1442 I_2 + 0.1375(OTFR) \tag{77}$$

where PTL and OTFR denote the percent of tank losses and opponent's tank force ratio, respectively for a given side.

In contrast to the results we found for Model 1 and Model 2, the negative coefficient of indicator variable I_1 indicates a Soviet advantage, and is significant. The positive coefficient of indicator variable I_2 indicates a defender advantage, and is also significant. The positive coefficient of the force ratio variable indicates that as the force

ratio increases, so do the force's losses. Again this is statistically significant but intuitively does not make much sense.

The fourth model uses the total aggregated tank force ratios as a predictor to predict the FER of tanks for each side. The FER of tanks for the blue forces is equal to the percent of blue tank losses divided by percent of red tank losses, and likewise for the FER of tanks for the red forces. Figures 58 and 59 show the FER of tanks plotted against Force ratio of tanks for Soviet and German forces, respectively. The representation of Model 4 looks like:

$$(BTL/BT)/(RTL/RT) = I_1 + I_2 + (BT/RT) \quad (78)$$

$$(RTL/RT)/(BTL/BT) = I_1 + I_2 + (RT/BT) \quad (79)$$

where I_1 indicates the difference between the blue force and red force, and I_2 indicates the difference between attacker and defender and have the values given in IV.B.7.(68) and IV.B.7.(69). BTL, RTL, BT and RT denote blue tank loss, red tank loss, the number of blue tanks and the number of red tanks, respectively.

Table 40 shows the coefficients, standard errors, and t values for Model 4.

	Value	Std. Error	t value	Pr(>|t|)
I1	-0.1242	1.8843	-0.0659	0.948
I2	2.2276	0.5959	3.7382	0.001
OTFR	0.2865	0.7517	0.3811	0.7064

Table 40. Important statistical values of the estimated coefficients of model 4.

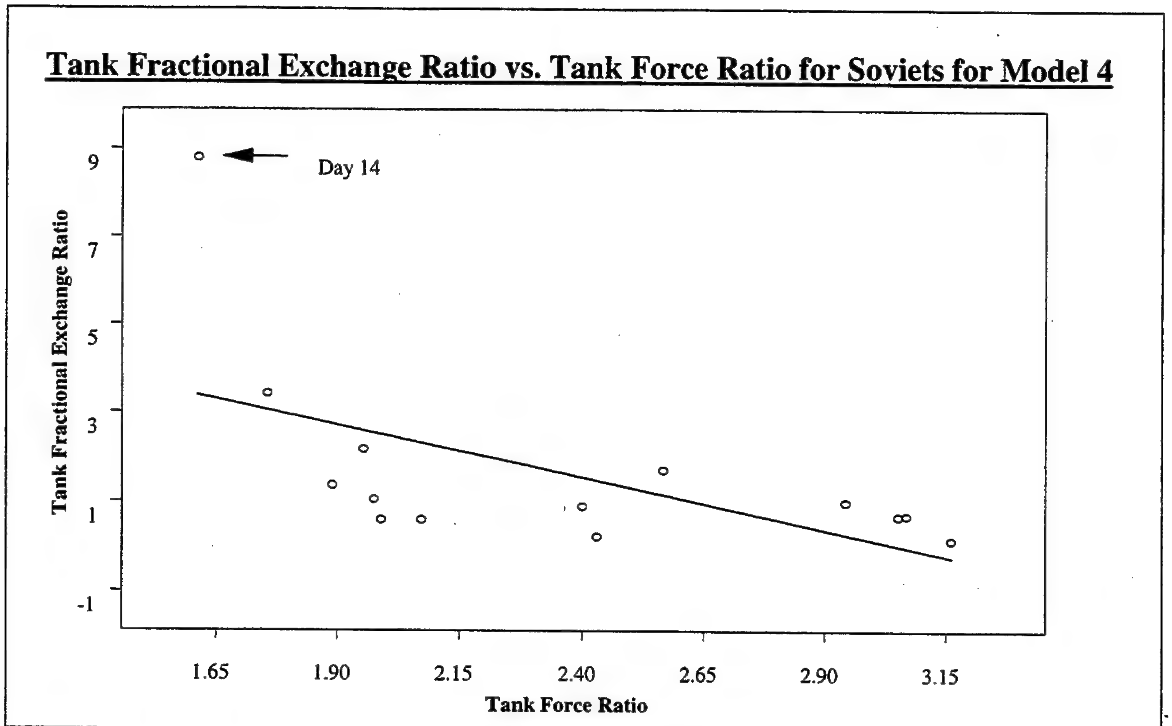

Figure 58. Tank fractional exchange ratio plotted versus tank force ratio for Soviet forces for model 4.

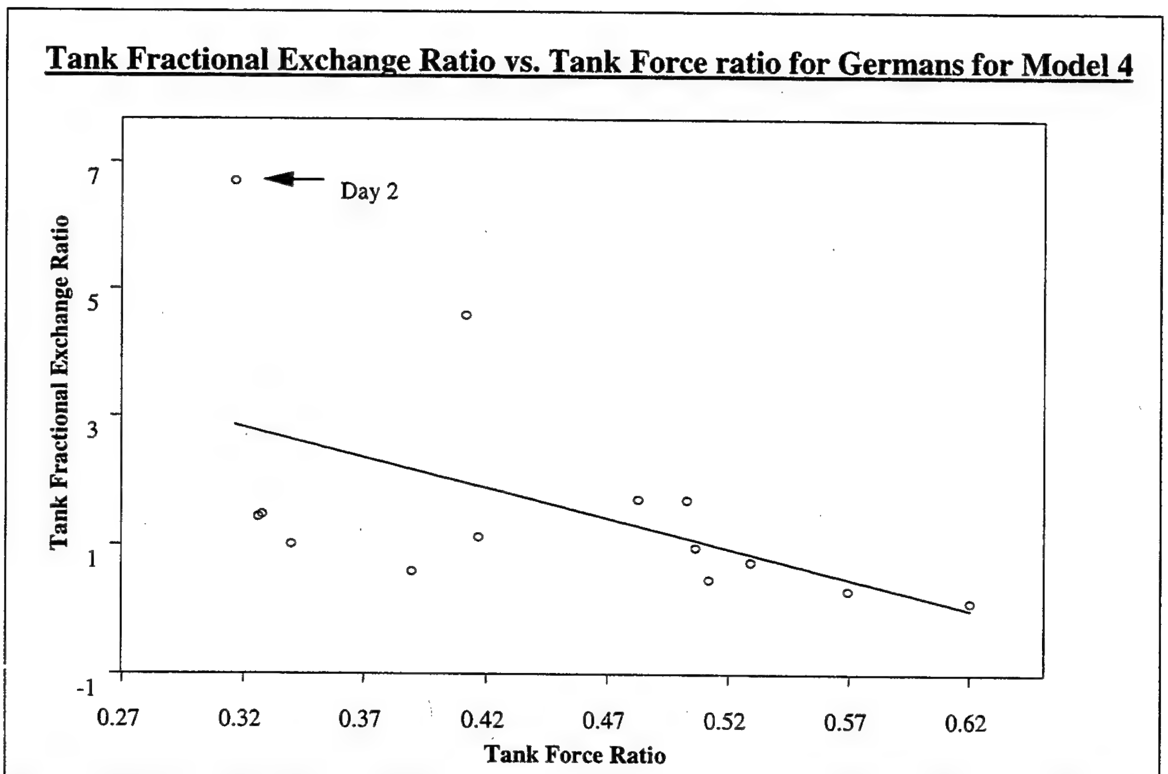

Figure 59.Tank fractional exchange ratio plotted versus tank force ratio for German forces for model 4.

The resulting model for Model 4, which gives an SSR value of 92.637 and a multiple R-squared value of 0.4941 is:

$$TFER = -0.1242 I_1 + 2.2276 I_2 + 0.2865(OTFR) \qquad (80)$$

where TFER and OTFR denote the FER of tanks and the opponent's tank FR for a given side.

Similar to the results we found for Model 3, the negative coefficient of indicator variable I_1, indicates a Soviet advantage and is not significant. The positive coefficient of indicator variable I_2 indicates a defender advantage and is significant. The positive coefficient of the force ratio variable indicates that as the force ratio increases so does your losses. Again the coefficient is not significant and intuitively does not make much sense.

The fifth model uses the same setup as Model 1, but it will do so by using the different weights first introduced in section IV.B.6 as the second weight combination, namely 1, 5, 15 and 20 for manpower, APC, artillery and tanks, respectively. Figures 60 and 61 show the loss ratio plotted against the force ratio for Soviet and German forces, respectively, using these weights. The presentation of model 5 looks like:

$$(\dot{B}/B) = I_1 + I_2 + (B/R) \qquad (81)$$

$$(\dot{R}/R) = I_1 + I_2 + (R/B) \qquad (82)$$

where I_1 indicates the difference between Blue force and Red force, and I_2 indicates the difference between attacker and defender and have the values given in IV.B.7.(68) and IV.B.7.(69).

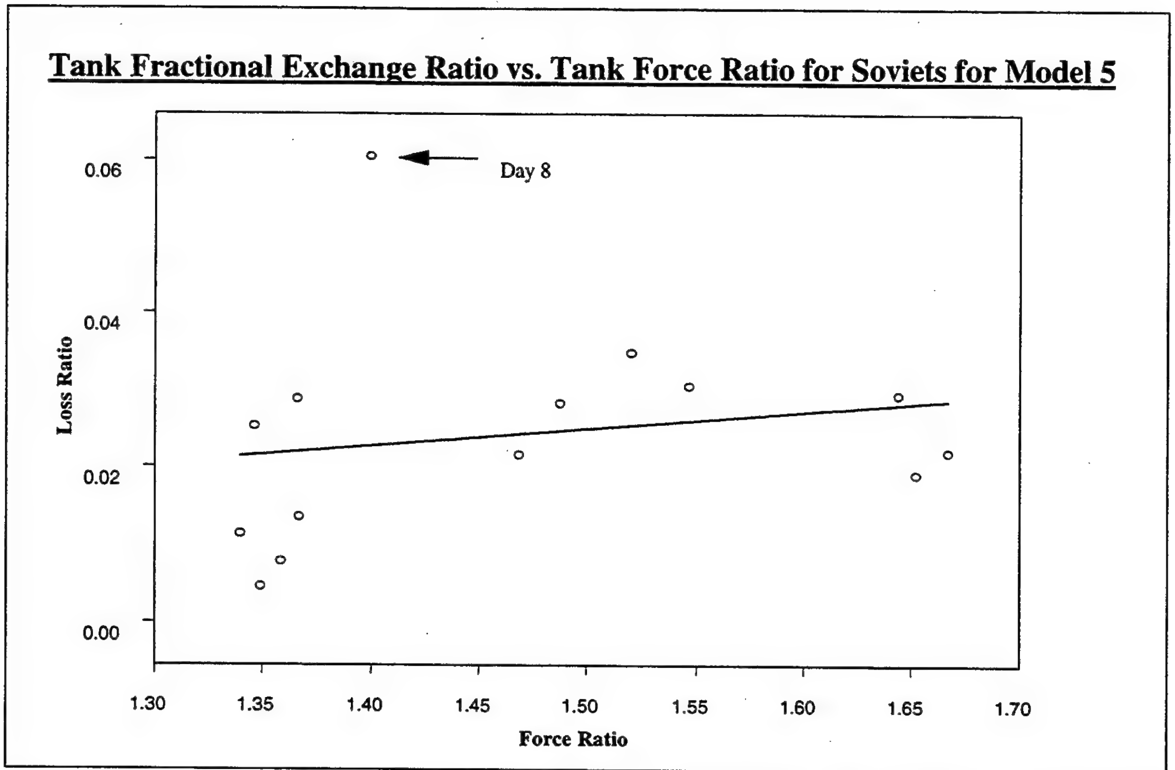

Figure 60. Loss ratio plotted versus force ratio for Soviet forces for model 5.

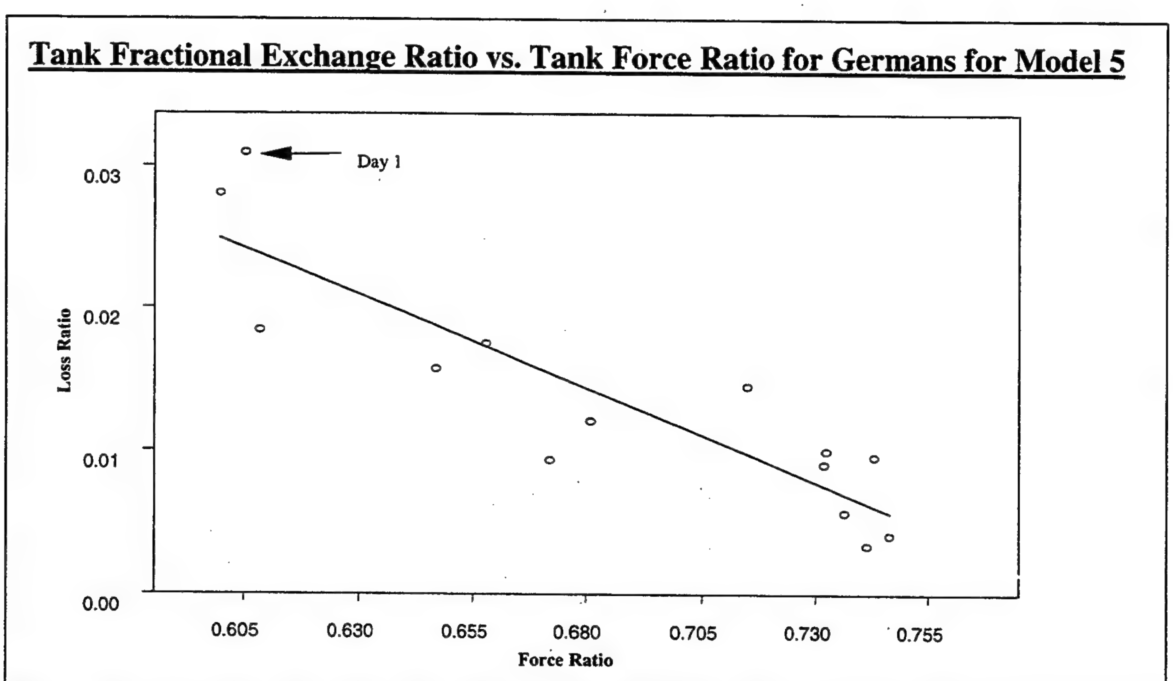

Figure 61. Loss ratio plotted versus force ratio for German forces for model 5.

where I_1 indicates the difference between Blue force and Red force, and I_2 indicates the difference between attacker and defender (time of battle) and have the values given in IV.B.7.(68) and IV.B.7.(69).

Table 41 shows the coefficients, standard errors, and t values for Model 5.

	Value	**Std. Error**	**t value**	**Pr(>|t|)**
I1	-0.0018	0.0072	-0:2544	0.8013
I2	0.0053	0.0039	1.3592	0.1862
OFR	0.0159	0.0049	3.2633	0.0032

Table 41. Important statistical values of the estimated coefficients of Model 5.

The resulting model for Model 5, which gives an SSR value of 0.0032546 and a multiple R-squared value of 0.7679 is:

$$PC = -0.0018 I_1 + 0.0053 I_2 + 0.0159(OFR) \tag{83}$$

where the notation has the same meaning as in Model 1.

Similar to the results we found for Model 3 and Model 4, the negative coefficient of indicator variable I_1 indicates a Soviet advantage and is not significant. The positive coefficient of indicator variable I_2 indicates a defender advantage and is not significant. The positive coefficient of the force ratio variable indicates that as the force ratio increases so do the losses. The coefficient is statistically significant and again, this interpretation intuitively does not make much sense.

In general, in the models we investigated in this section, the indicator variable I_2 is always positive, different from the result we found in the sections, which investigated

the tactical parameter *d*. This observation suggests that it is advantageous to be the defender, not the attacker. Another interesting, yet ironic, result is the positive force ratio coefficient found in the models throughout the section, suggesting that the more powerful you are, the more you lose, which intuitively does not make much sense.

When the plots are investigated it is seen that, the higher the force ratio or FER is, the less the loss is, except for the Soviets in Model 1, Model 3 and Model 5. So, the results are telling somewhat different than what the plots are telling. This may be due to the interpretation that fitting the logarithmically transformed equations does not necessarily gives the best fit in the original form.

Table 42 summarizes the results found in this section.

	I1	I2	Predictor	Multiple R-squared
Model 1	0.001	0.0048	0.0147	0.7699
Model 2	1.1849	0.5647	0.4153	0.6963
Model 3	-0.2703	0.1442	0.1375	0.7077
Model 4	-0.1242	2.2276	0.2865	0.4941
Model 5	-0.0018	0.0053	0.0159	0.7679

Table 42. Results for the section investigating the force ratio and the fractional exchange ratio models.

When the overall results given in Table 42 are examined it is seen that Model 1 and Model 2 have positive I_1 coefficients, which indicates a German advantage while the rest of the models have negative I_1 coefficients, which indicates a Soviet advantage. All

models have positive I_2 coefficients, which indicates a defender advantage. The first model with the highest multiple R-squared value gives the best fit.

8. Fitting the standard Lanchester equations

This section fits the basic Lanchester Equations, (i.e., Lanchester Linear, Lanchester Square and Lanchester Logarithmic models), to the Battle of Kursk data. The basic Lanchester equations are given in I.B.(1) and I.B.(2).

For the Lanchester linear model where p=q=1, the loss for one side will be equal to the product of the existing number of forces of both sides, and a coefficient. The Lanchester linear model will look like;

$$\dot{B} = aRB \tag{84}$$

$$\dot{R} = bBR \tag{85}$$

This model is solved like a typical regression through the origin equation and the resulting model for the Lanchester linear model, which gives an SSR value of 6.24×10^{8} is:

$$\dot{B} = 6.6834\times10^{-8}RB \tag{86}$$

$$\dot{R} = 2.6893\times10^{-8}BR \tag{87}$$

Figures 62 and 63 show fitted losses plotted versus real losses for Soviet and German forces, respectively, for the Lanchester linear model.

For the Lanchester Square Model, where p=1 and q=0, the loss for one side will be equal to the product of the existing number of forces of the opponent and a coefficient. The Lanchester square model will look like;

$$\dot{B} = aR \tag{88}$$

$$\dot{R} = bB \tag{89}$$

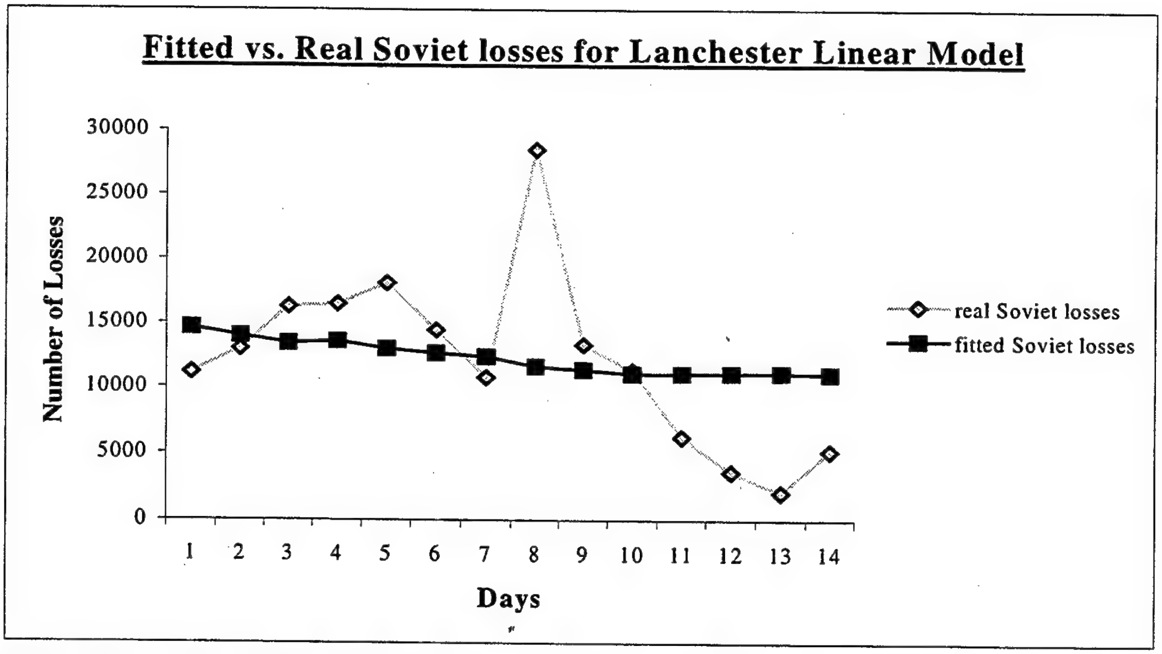

Figure 62. Fitted losses plotted versus real losses for Soviet forces for the Lanchester linear model. The same three-phase pattern where the model over/underestimates the battle in three distinctive phases is observable in this plot for the model, which uses the Lanhester linear model, too.

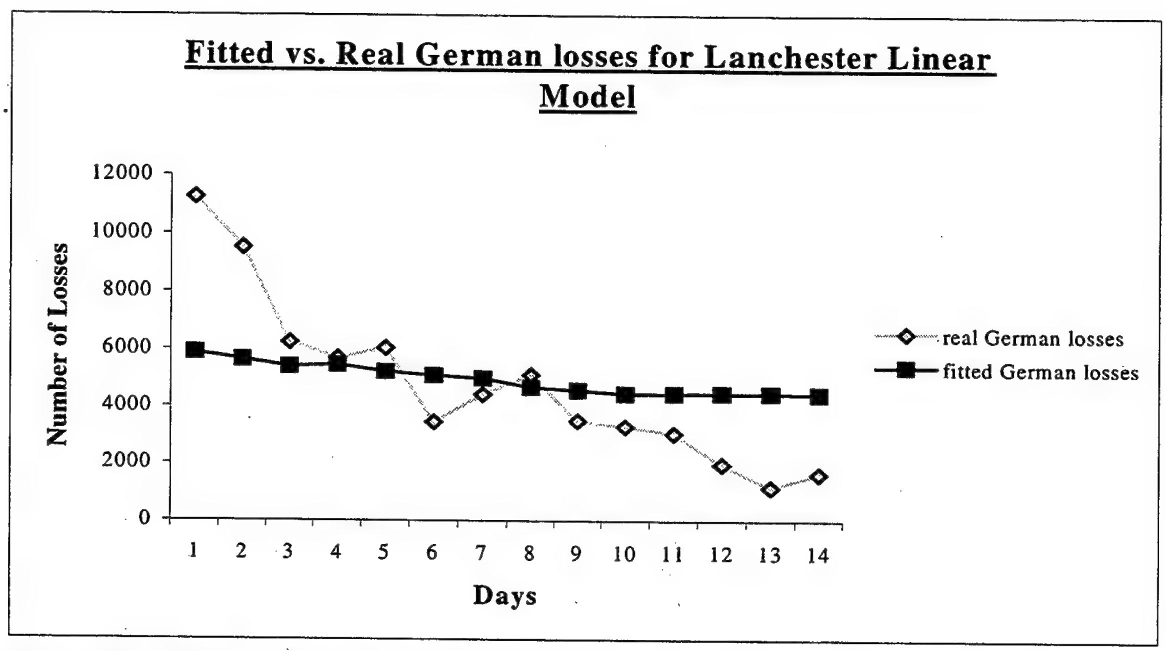

Figure 63. Fitted losses plotted versus real losses for German forces for the Lanchester linear model. Eight days are underestimated while six days are overestimated.

The resulting model for Lanchester square model that gives an SSR value of 6.79×10^{8} is:

$$\dot{B} = 0.0335R \tag{90}$$

$$\dot{R} = 0.0098B \tag{91}$$

The high value of the a parameter in the above equation indicates that the Germans fought three times better than the Soviets.

Figures 64 and 65 show fitted losses plotted versus real losses for Soviet and German forces, respectively, for the Lanchester square model.

For the Lanchester logarithmic model where p=0 and q=1, the loss for one side will be equal to the product of the existing number of forces of its own and a coefficient. Lanchester logarithmic model will look like:

$$\dot{B} = aB \tag{92}$$

$$\dot{R} = bR \tag{93}$$

The resulting model for Lanchester logarithmic model, which gives an SSR value of 6.57×10^{8} is:

$$\dot{B} = 0.0243B \tag{94}$$

$$\dot{R} = 0.0131R \tag{95}$$

Figures 66 and 67 show fitted losses plotted versus real losses for Soviet and German forces, respectively, for the Lanchester logarithmic model.

The basic Lanchester Equations do not give the best fit for the Battle of Kursk data. Out of the three Lanchester Models analyzed, the Lanchester linear model gives the best fit (i.e., smallest SSR value).

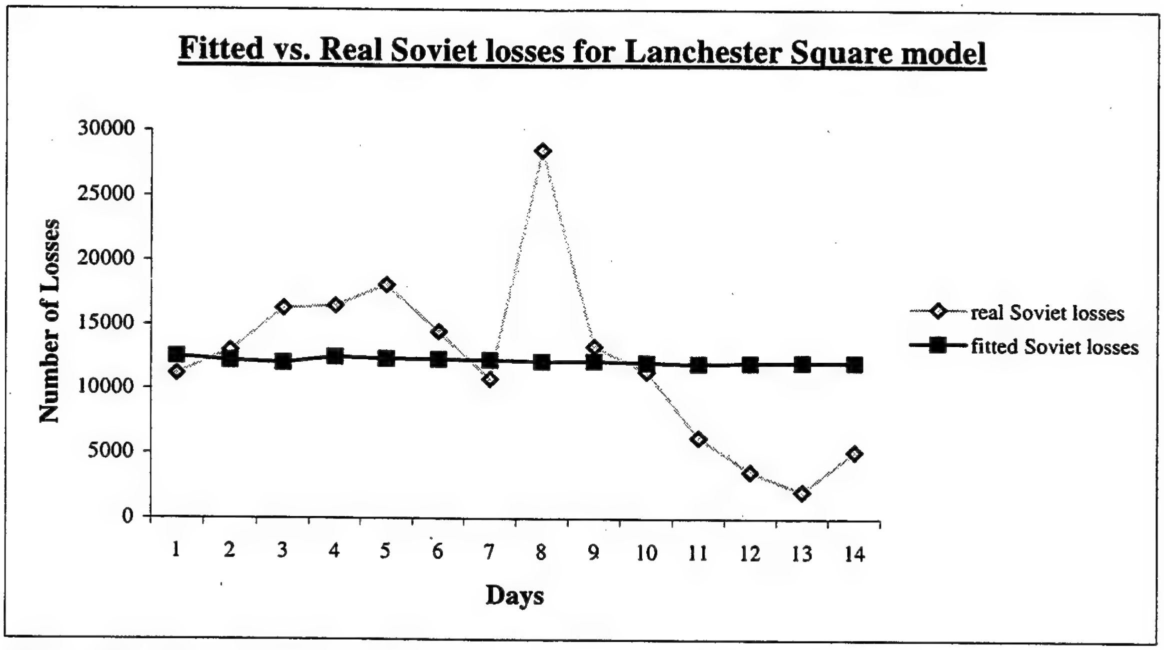

Figure 64. Fitted losses plotted versus real losses for Soviet forces for the Lanchester square model. The same three-phase pattern where the model over/underestimates the battle in three distinctive phases is observable in this plot for the model, which uses the Lanhester square model, too.

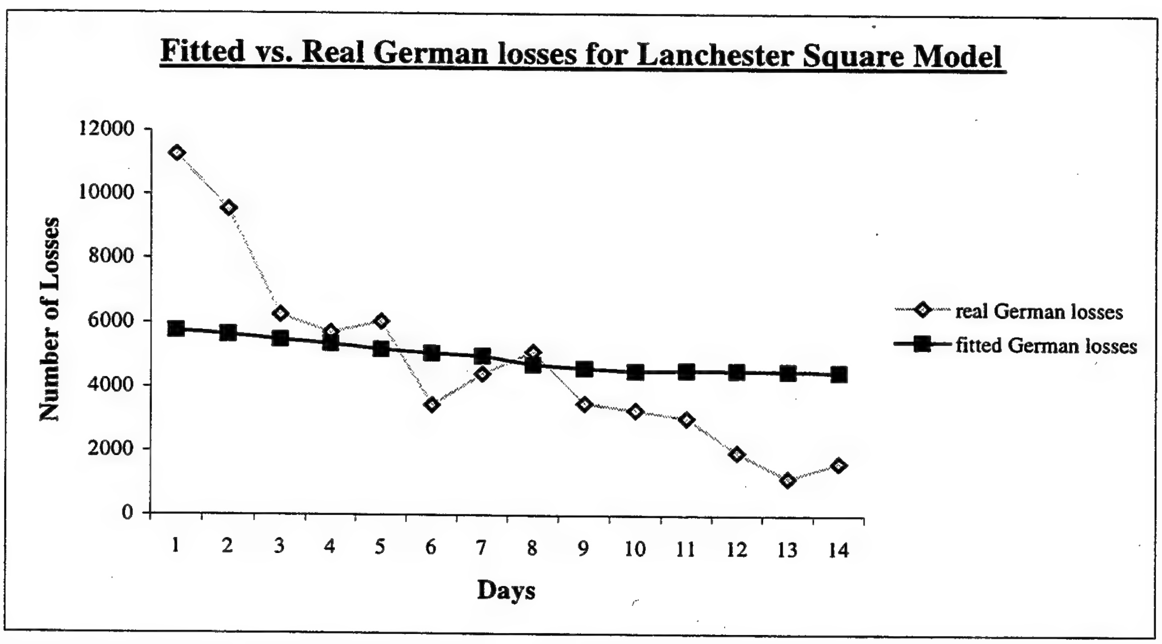

Figure 65. Fitted losses plotted versus real losses for German forces for the Lanchester square model. Eight days are underestimated while six days are overestimated.

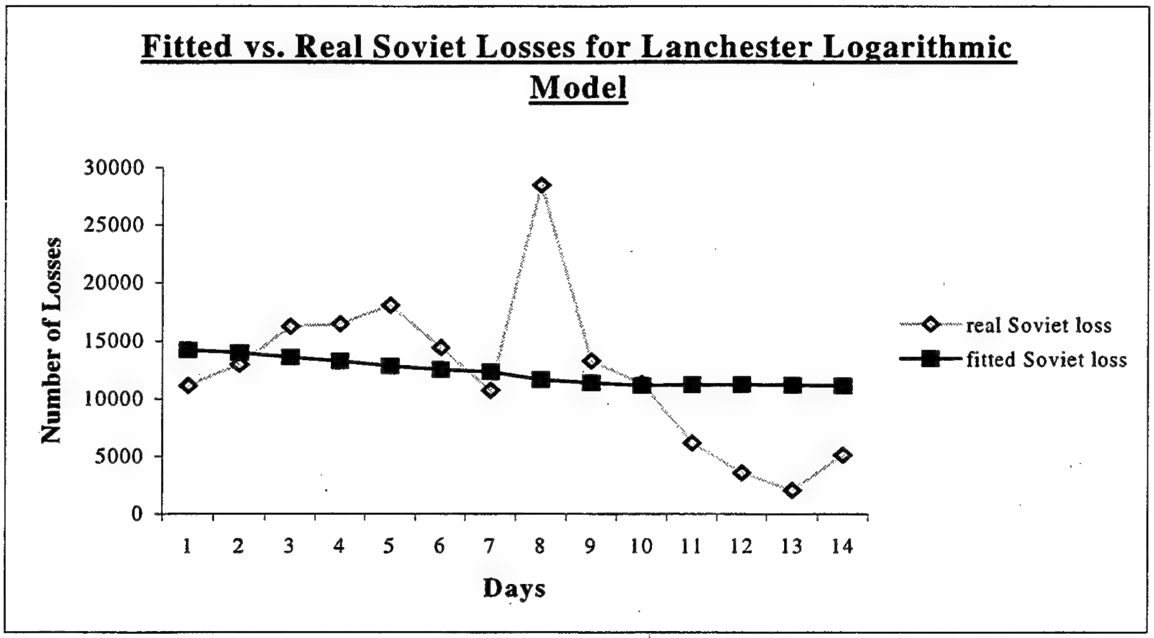

Figure 66. Fitted losses plotted versus real losses for Soviet forces for the Lanchester logarithmic model.

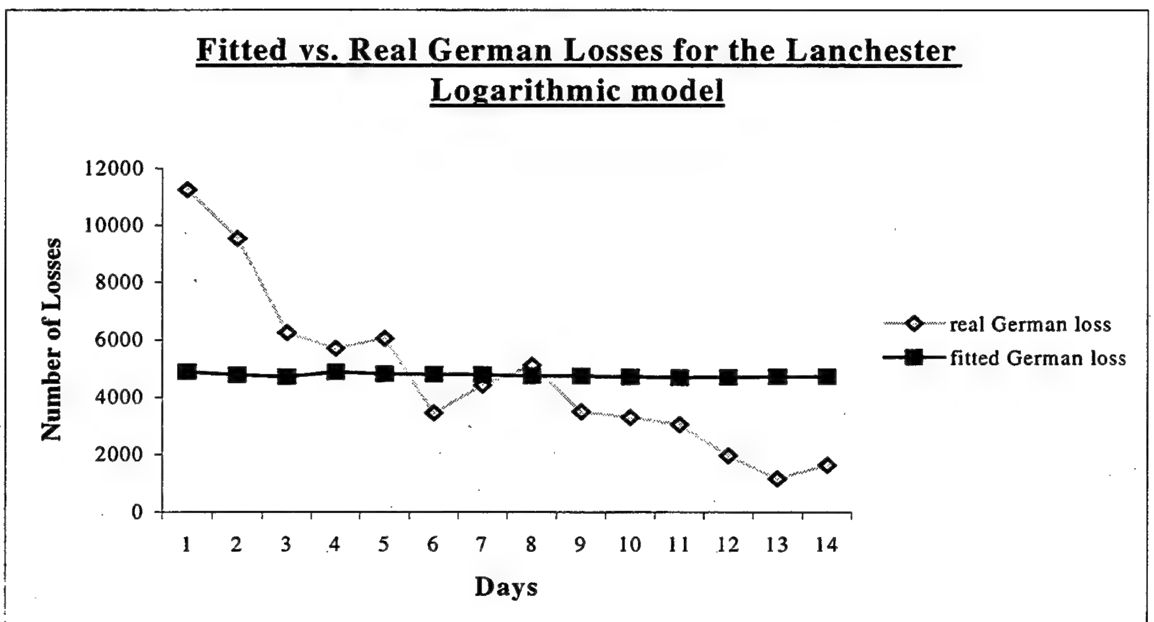

Figure 67. Fitted losses plotted versus real losses for German forces for the Lanchester logarithmic model.

Again in all Lanchester Models, the *a* and *b* parameters are significantly small and $a>b$.

Fricker's findings were closest to Lanchester's logarithmic model, while Bracken's findings were closest to Lanchester's linear model. Out of the three basic Lanchester models, it is the Lanchester linear model that best fits the Battle of Kursk data. The Lanchester logarithmic model gives the second best fit for the Battle of Kursk data, while the Lanchester square model gives the third best (i.e., the worst) fit for the Battle of Kursk data.

9. Fitting Morse-Kimball equations

This section will fit the Morse-Kimball Equations to the Battle of Kursk data. Morse and Kimball suggest that one side's losses do not depend solely on the opponent's forces, losses also depend on one's own failures and other mechanical breakdowns too, like the case in the logarithmic law. The Morse-Kimball Equations are:

$$\dot{B} = aR + \alpha_1 B \tag{96}$$

$$\dot{R} = bB + \alpha_2 R \tag{97}$$

These equations are fit separately for the Germans and the Soviets, and the resulting model for the Morse-Kimball Equations, which gives an SSR value of 5.51×10^8 and an R^2 value of 0.2297 is:

$$\dot{B} = -0.0412R + 0.0537B \tag{98}$$

$$\dot{R} = 0.0603B - 0.0707R \tag{99}$$

Figures 68 and 69 show fitted losses plotted versus real losses for Soviet and German forces, respectively, for the Morse-Kimball Equations model.

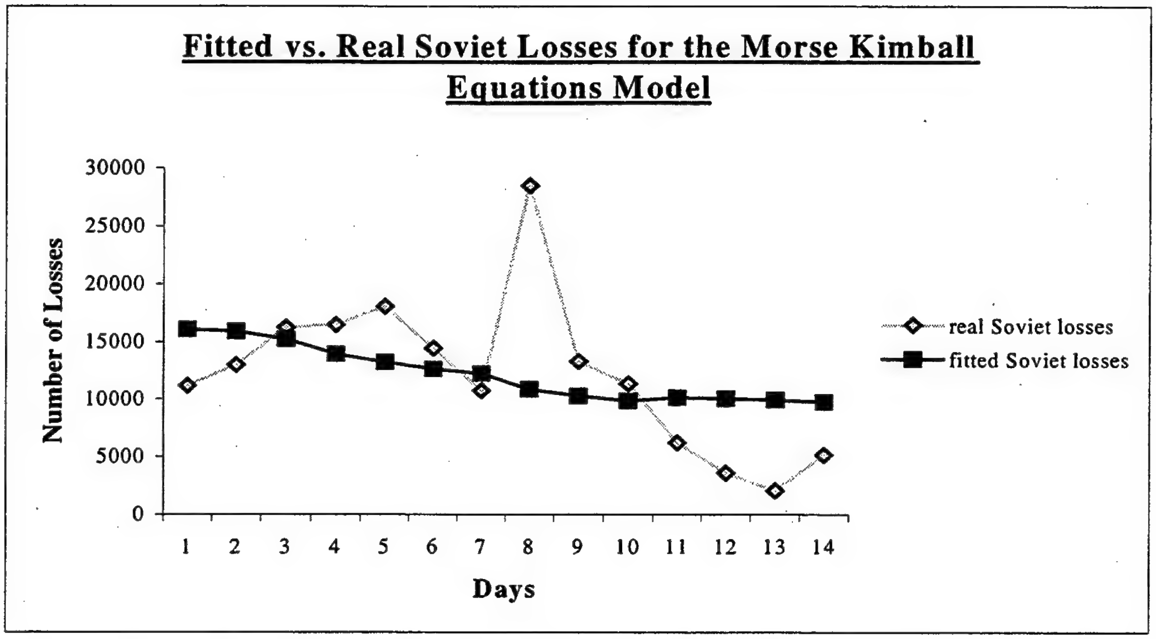

Figure 68. Fitted losses plotted versus real losses for the Soviet forces for the model using Morse Kimball Equations. The same three-phase pattern where the model over/underestimates the battle in three distinctive phases is observable in this plot for the model, which uses Morse Kimball equations, too.

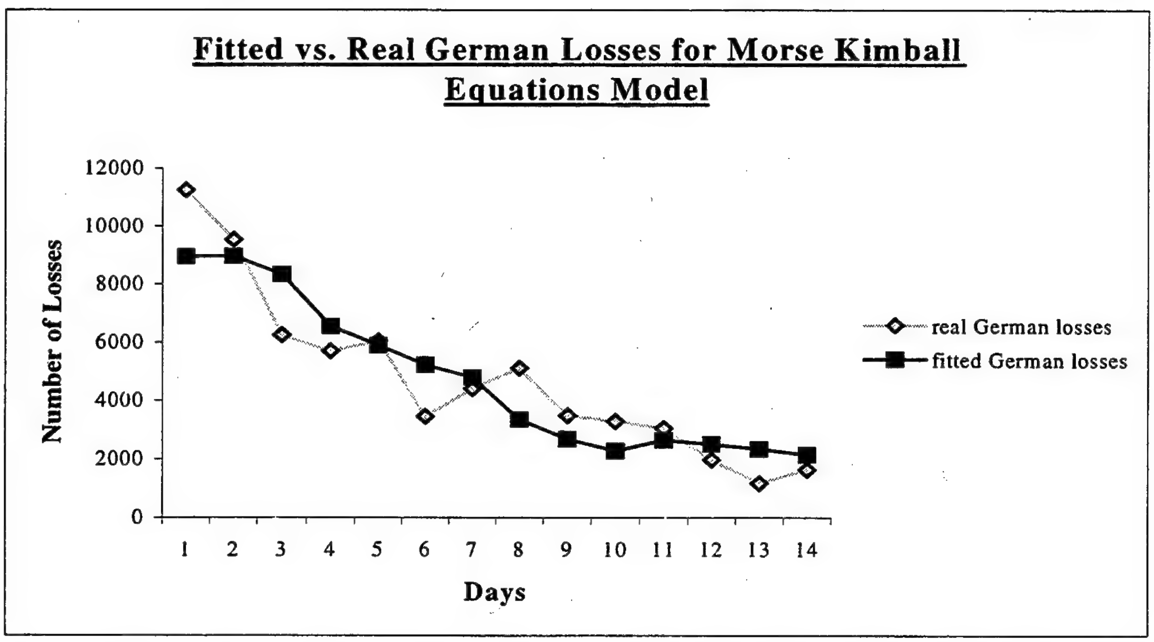

Figure 69. Fitted losses plotted versus real losses for the German forces for the Morse Kimball Equations Model.

Fitting Morse-Kimball Equations to the Battle of Kursk data improves the fit. The SSR value of 5.51×10^{8} is one of the lowest SSR values we have so far. But, just as in the models used for the change points approach for each side in section IV.B.5, the parameters physically do not make sense.

For the blue force, the negative a parameter indicates that the more the red forces there are, the less the number of blue casualties. For the red force, the negative α_2 parameter indicates that the greater the number of the red forces is, fewer red casualties are going to be. This physically does not make much sense; so, even if fitting Morse-Kimball equations give a low SSR value of 5.51×10^{8}, we cannot accept this fit.

10. Fitting the parameters found by Bracken and Fricker

In this section, the parameters for the Ardennes data found in Bracken and Fricker's studies will be used to fit the Battle of Kursk data.

a. Bracken's parameters

In his study, Bracken's conclusion for the Lanchester Model with the tactical parameter is given as:

$$\dot{B} = 8\times10^{-9}(\frac{8}{10} \text{ or } \frac{10}{8})R^{1}B^{1} \tag{100}$$

$$\dot{R} = 1\times10^{-8}(\frac{10}{8} \text{ or } \frac{8}{10})B^{1}R^{1} \tag{101}$$

Figures 70 and 71 show fitted losses plotted versus real losses for Soviet and German forces, respectively, for the parameters of Bracken's model (with the tactical parameter) given above, which yields an SSR value of 2.39×10^{9} for the Battle of Kursk data.

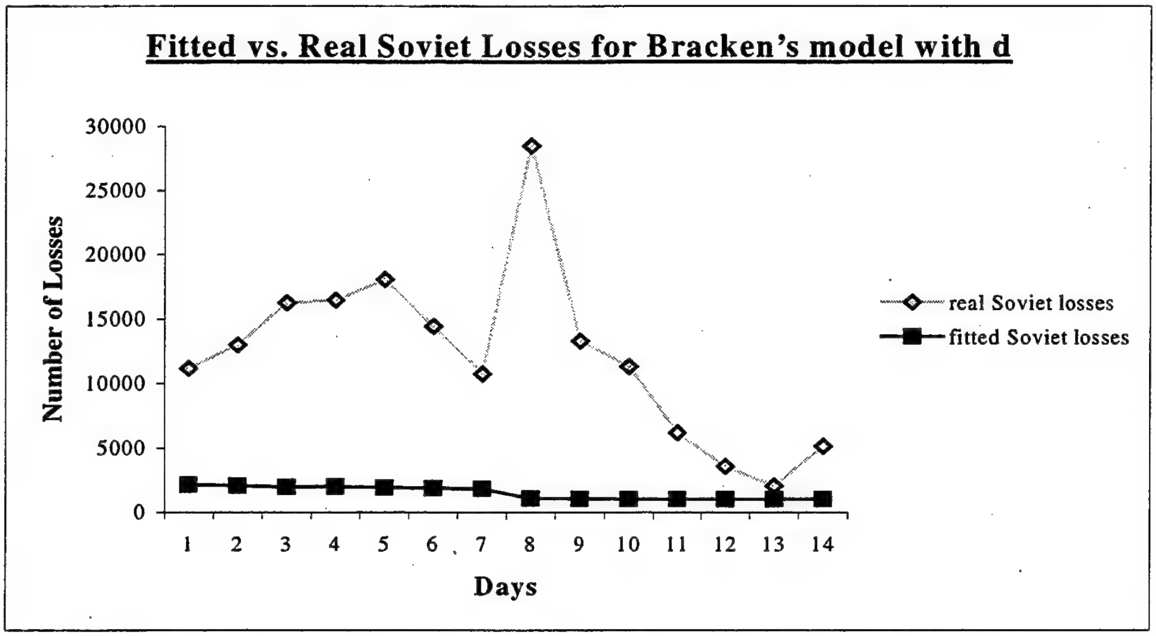

Figure 70. Fitted losses plotted versus real losses for Soviet forces for Bracken's model with the tactical parameter *d*. Bracken's Ardennes parameters always underestimated the Soviet losses for the Battle of Kursk.

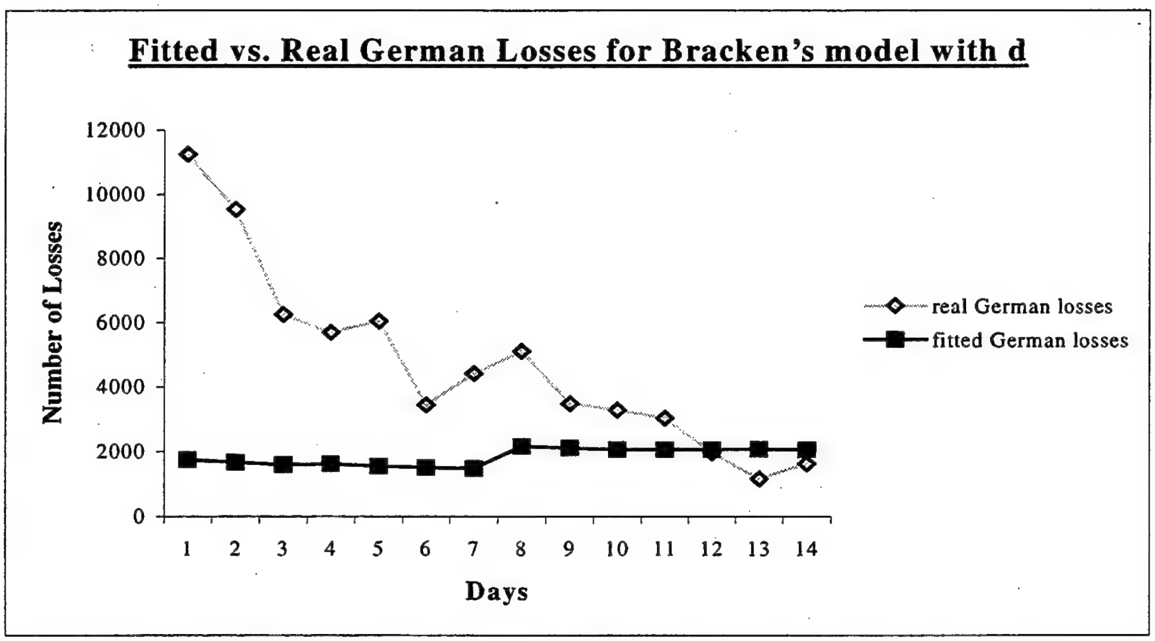

Figure 71. Fitted losses plotted versus real losses for German forces for Bracken's model with the tactical parameter *d*. Except the last three days of the battle, Bracken's Ardennes parameters always underestimated the German losses for the whole Battle of Kursk data.

Bracken's conclusion for the Lanchester model without the tactical parameter is given as:

$$\dot{B} = 8\times10^{-9} R^{1.3} B^{0.7} \tag{102}$$

$$\dot{R} = 1\times10^{-8} B^{1.3} R^{0.7} \tag{103}$$

Figures 72 and 73 show fitted losses plotted versus real losses for Soviet and German forces, respectively, for the parameters of Bracken's model (without the tactical parameter) given above, which yields an SSR value of 2.46×10^{9} for the Battle of Kursk data.

Fitting Brackens's parameters to the Battle of Kursk data does not improve the model's fit and gives the highest SSR value thus far. It is significant that Bracken's parameters always underestimates the real casualties for the Battle of Kursk data.

b. Fricker's parameters

In Fricker's study, the conclusion for the Lanchester model with the tactical parameter is given as:

$$\dot{B} = 4.7\times10^{-27}(\frac{1}{0.8093} \text{ or } 0.8093)B^{5} \tag{104}$$

$$\dot{R} = 3.1\times10^{-26}(0.8093 \text{ or } \frac{1}{0.8093})R^{5} \tag{104}$$

Figures 74 and 75 show fitted losses plotted versus real losses for Soviet and German forces, respectively, for the parameters of Fricker's model (with the tactical parameter) given above that yields an SSR value of 3.02×10^{9} for the Battle of Kursk data.

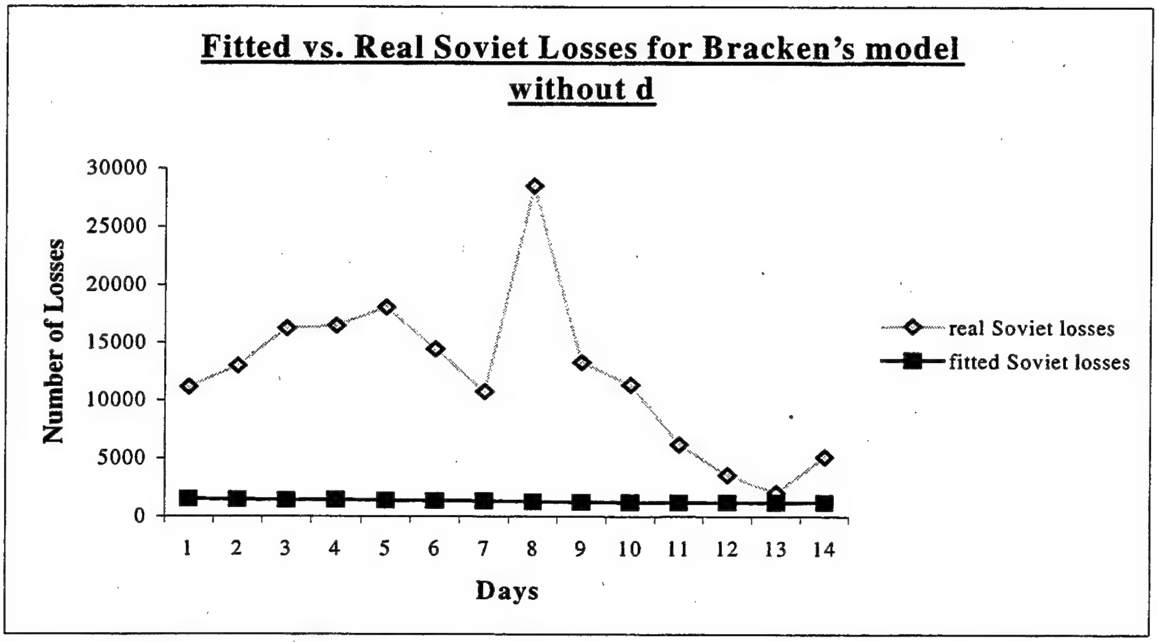

Figure 72. Fitted losses plotted versus real losses for the Soviet forces for Bracken's model without the tactical parameter *d*. Bracken's Ardennes parameters always underestimated the Soviet losses for the Battle of Kursk data.

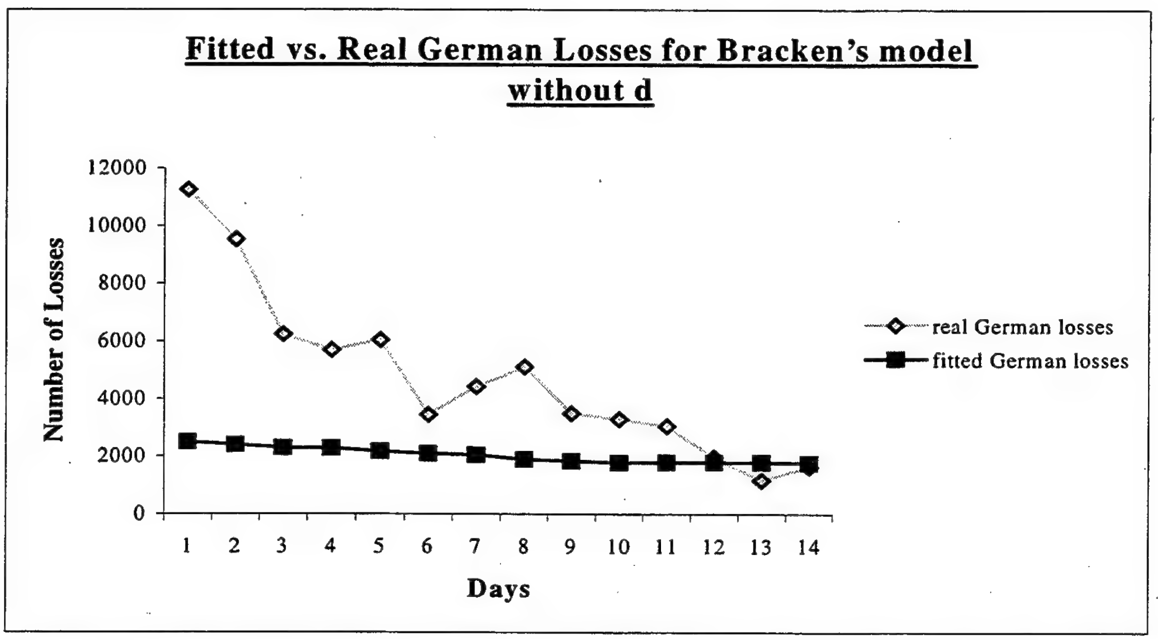

Figure 73. Fitted losses plotted versus real losses for the German forces for Bracken's model without the tactical parameter *d*. With the exception of the last three days of the battle, Bracken's Ardennes parameters always underestimated the German losses for the whole Battle of Kursk data.

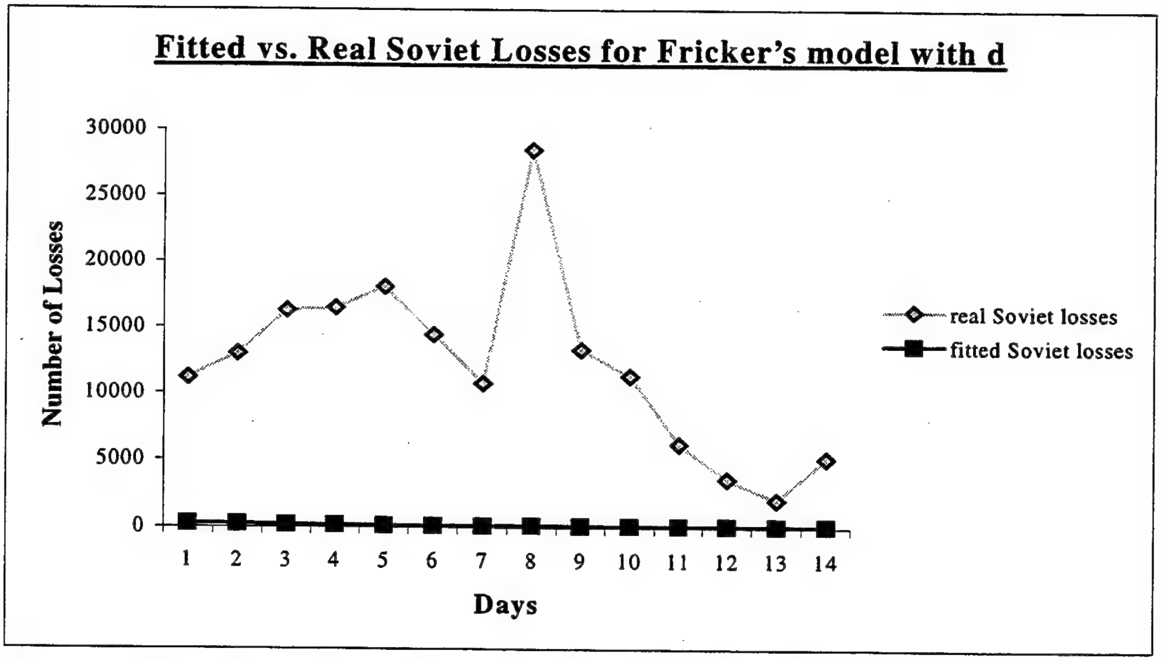

Figure 74. Fitted losses plotted versus real losses for Soviet forces for Fricker's model with the tactical parameter *d*. Fricker's Ardennes parameters always underestimated the Soviet losses for the Battle of Kursk.

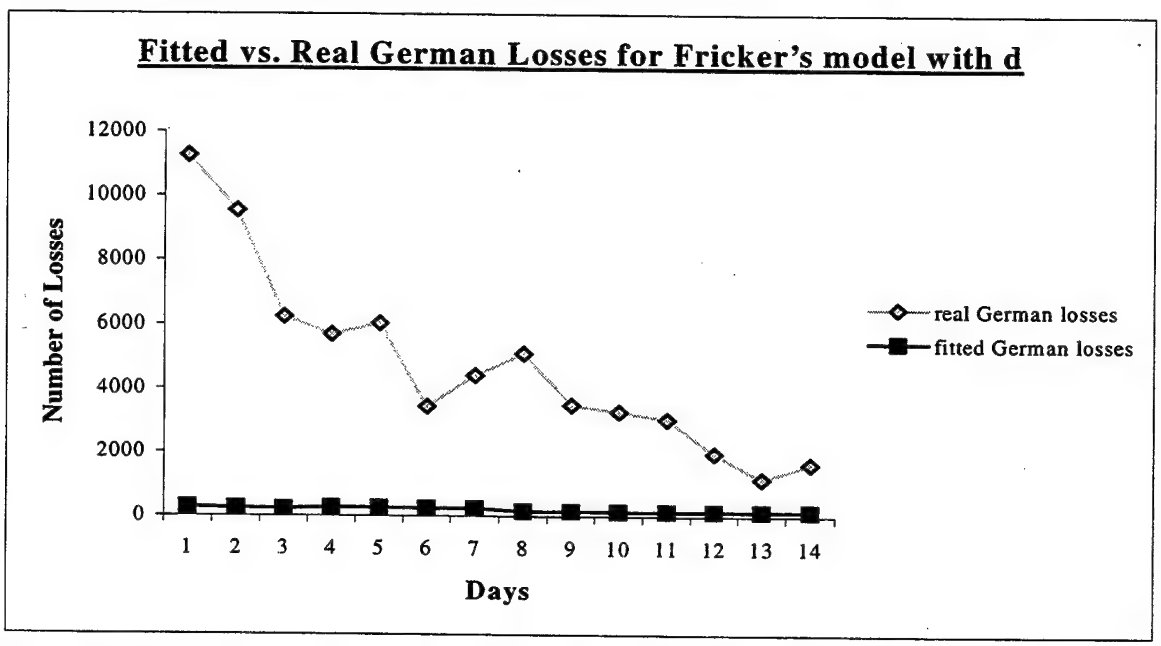

Figure 75. Fitted losses plotted versus real losses for Soviet forces for Fricker's model with the tactical parameter *d*. Fricker's Ardennes parameters always underestimated the German losses for the Battle of Kursk.

Fricker's conclusion for the Lanchester model with the air sortie data added is given as:

$$\dot{B} = 2.7\times10^{-24}(\frac{1}{0.7971} \text{ or } 0.7971)B^{4.6} \tag{106}$$

$$\dot{R} = 1.6\times10^{-23}(0.8093 \text{ or } \frac{1}{0.8093})R^{4.6} \tag{107}$$

Figures 76 and 77 show fitted losses plotted versus real losses for Soviet and German forces, respectively, for the parameters of Fricker's model (with air sortie data added) given above, which yields an SSR value of 2.77×10^{9} for the Battle of Kursk data.

Like Bracken's models, fitting Fricker's parameters to the Battle of Kursk data does not improve the model's fit, it gives the highest SSR value in this study so far. Fricker's parameters always underestimate the real casualties for the Battle of Kursk data. This finding is similar to the one for Bracken's parameters.

In general, fitting Bracken's or Fricker's Ardennes parameters to the Battle of Kursk data does not improve the fit; they both give the highest SSR value we have in this study so far. This result suggests that the parameters of one battle data cannot be used to predict another. Each battle has its own unique parameters which cannot be applied to another one battle.

Another interesting finding is that when Bracken's or Fricker's Ardennes parameters are applied to Kursk data, they always underestimate the daily attrition rates. This finding suggests that Battle of Kursk was a much more intense battle than the Ardennes campaign.

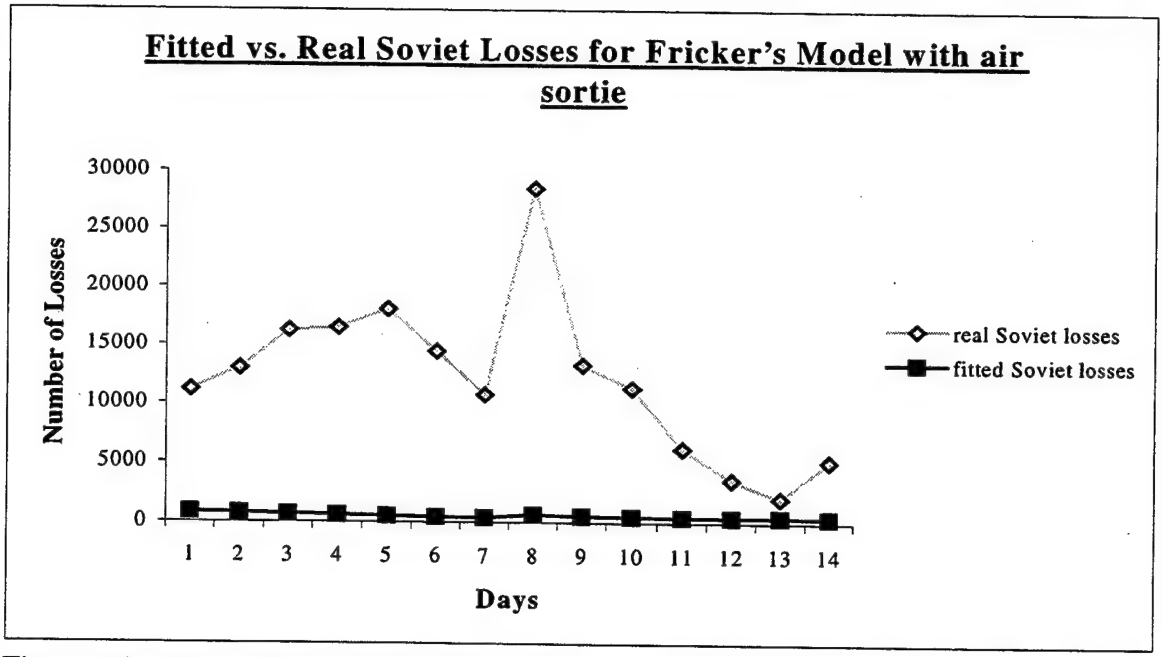

Figure 76. Fitted losses plotted versus real losses for Soviet forces for Fricker's model with the air sortie data added. Notice that Fricker's Ardennes parameters always underestimated the Soviet losses for the Battle of Kursk.

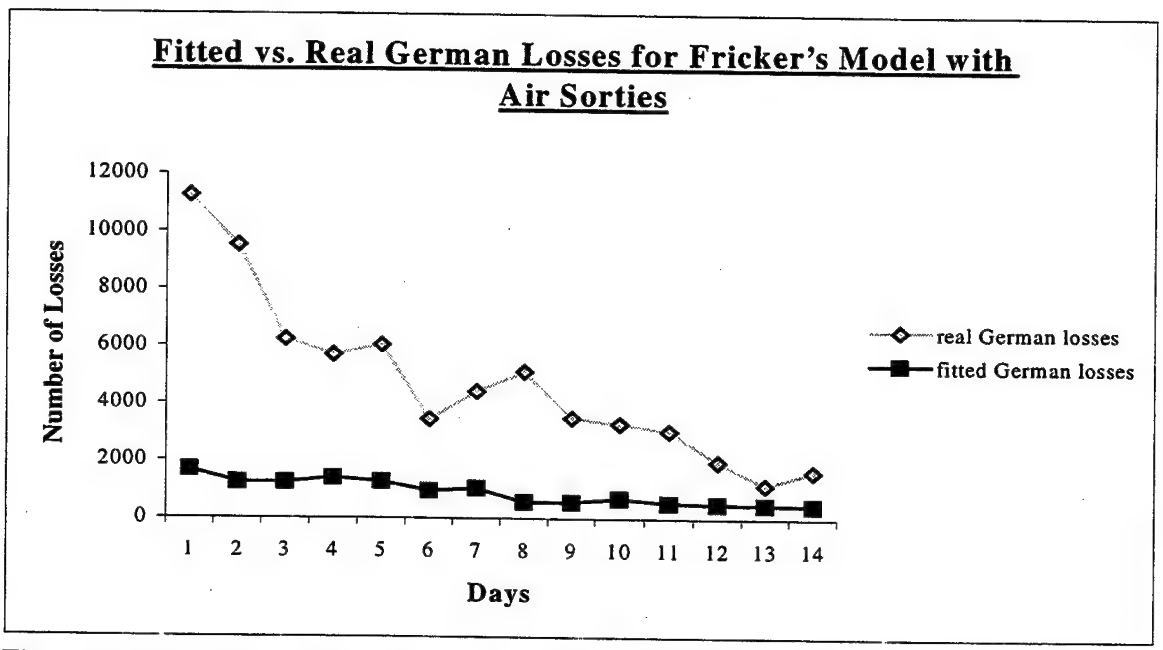

Figure 77. Fitted losses plotted versus real losses for German forces for Fricker's model with air sortie data added. Notice that Fricker's Ardennes parameters always underestimated the German losses for the Battle of Kursk.

11. Summary of results

This section summarizes the results of all the models explored in previous chapters.

Name of the model	*a*	*b*	*p*	*q*	*d*	*SSR*	R^2
Bracken Model 1 Ardennes	8.0E-9	1.0E-8	1.0	1.0	1.25	1.63E+9	0.2552
Bracken Model 2 Ardennes	8.0E-9	1.0E-8	1.3	0.7	1.0	2.08E+8	0.0493
Bracken Model 1 Kursk	1.2E-8	9.0E-9	0.1	2.0	0.9	8.65E+8	0.0006
Bracken Model 3 Kursk	1.2E-8	9.0E-9	0.3	1.8	1.0	8.88E+8	-0.0266
Frick.Ard. w/o sorties with d	4.7E-27	3.1E-26	0	5	0.8093	1.57E+8	-0.7938
Frick.Ard. w sorties with d	2.7E-24	1.6E-23	0	4.6	0.7971	2.64E+7	0.5256
Frick.Kursk w/o sorties with d	3.76E-33	1.09E-32	0.0604	6.3066	0.79	5.94E+8	0.1703
Frick.Kursk w/o sorties w/o d	1.61E-33	3.44E-33	3.6736	2.6934	-	2.16E+9	0.0657
Frick.Kursk with sorties with d	3.35E-27	5.76E-27	0.0955	5.2207	0.93	6.23E+8	0.1294
Frick.Kursk with sorties w/o d	5.01E-27	3.85E-27	1.4983	3.8179	-	7.16E+8	-0.0222
Clemens Linear Regression	6.92E-49	6.94E-48	5.3157	3.6339	-	1.13E+9	0.9975
Clemens Newton-Raphson	3.73E-6	5.91E-6	0.0	1.6178	-	1.04E+9	-0.6242
Linear Regression Model	1.06E-47	1.90E-48	5.7475	3.3356	-	6.36E+8	0.1126
Robust LTS Regression	2.27E-40	1.84E-41	6.0843	1.7312	-	5.54E+8	0.2262

Name of the model	*a*	*b*	*p*	*q*	*d*	*SSR*	R^2
Lin.Reg. With Air sorties	1.40E-30	2.09E-36	5.1323	1.7793	-	6.85E+8	0.0433
Robust LTS with Air sorties	1.21E-38	1.75E-39	5.3691	2.0883	-	7.58E+8	-0.0579
Linear Regression With *d*	1.88E-47	1.07E-48	7.5038	1.5793	1.17	6.24E+8	0.1295
Robust LTS With *d*	2.27E-40	1.84E-41	6.0843	1.7312	1.0	5.54E+8	0.2262
Campaign in four Parts	1.88E-47	1.07E-48	7.5038	1.5793	4 periods *d*=0.91,1.24, 1.0,1.17	5.34E+8	-2.3410
Campaign in four Parts	1.88E-47	1.07E-48	7.5068	1.5793	4 periods *d*=0.91,1.24, 0.32,1.17	1.69E+8	-0.0607
Campaign in four Parts	1.88E-47	1.07E-48	7.5038	1.5793	1.14	1.89E+8	0.5689
Campaign in four Parts	1.85E-51	3.56E-53	9.6853	0.1458	-	1.90E+8	0.5658
Change Point 7/7	8.91E-30	2.62E-31	6.4117	-0.4323	-	1.53E+8	0.7448
Change Point 7/7	1.90E-232	4.37E-291	18.0587	34.4502	-	1.53E+8	0.7448
Change Point 8/6	7.75E-5	1.91E-6	4.4212	-2.8454	-	2.43E+8	0.3488
Change Point 8/6	1.94E-246	1.32E-247	25.7652	18.7674	-	2.43+E8	0.3488
Weight comb.1 Lin.Reg.	1.25E-38	1.60E-39	5.2298	2.2746	-	1.15E+9	0.0870
Weight Comb.1 Rob.LTS	7.26E-35	5.53E-36	5.5312	1.3268	-	1.07E+9	0.1514
Weight comb.2 Lin.Reg.	2.50E-46	3.49E-47	5.7638	3.1222	-	6.24E+8	0.0975
Weight Comb.2 Rob.LTS	7.85E-36	4.75E-37	5.8613	1.1899	-	5.48E+8	0.2072
Weight comb.3 Lin.Reg.	3.78E-39	5.34E-40	5.2293	2.3513	-	1.15E+9	0.0926

Name of the model	a	b	p	q	d	SSR	R^2
Weight Comb.3 Rob.LTS	1.46E-35	9.33E-37	5.9619	1.0159	-	1.06E+9	0.1637
Weight comb.4 Lin.Reg.	2.89E-42	3.91E-43	5.4863	2.6660	-	8.63E+9	0.0943
Weight Comb.4 Rob.LTS	5.05E-35	3.51E-36	5.6294	1.2631	-	7.74E+8	0.1873
Lanchester Linear model	6.68E-8	2.68E-8	1.0	1.0	-	6.24E+8	0.1290
Lanchester Square model	0.0335	0.0098	1.0	0	-	6.79E+8	0.0521
Lanchester Logarithmic model	0.0243	0.0131	0	1.0	-	6.57E+8	0.0831
Morse Kimball Equations	a=-0.041	α_1=0.053	b=0.060	α_2=-0.07	-	5.51E+8	0.2297
Bracken's Parameters with d	8.0E-9	1.0E-8	1.0	1.0	1.25	2.39E+9	-2.4235
Bracken's Parameters w/o d	8.0E-9	1.0E-8	1.3	0.7	-	2.46E+9	-2.4430
Fricker's Parameters with d	4.7E-27	3.1E-26	0	5.0	0.8093	3.02E+9	-3.2123
Fricker's Par.s with air sortie	2.7E-24	1.6E-23	0	4.6	0.7971	2.79E+9	-2.9021

Table 42. Results of all the models explored and investigated in Chapter IV.

The R^2 value (0.9975) given for Clemens' linear regression model is the self reported value by Clemens and must have been calculated differently than the R^2 values calculated throughout the thesis. When recomputed, a negative R^2 value is found.

Clemens provided four digits of precision in his estimates of p and q, while Bracken and Fricker gave two. The R^2 values that are found for the models, which do not use the parameters from other studies, are calculated using parameters with four digits

of precision. The slightly negative R^2 values found for some of these models are not a result of using low precision.

When the above results are examined, it is seen that the best fitting model for the Battle of Kursk data is the robust LTS regression model used in section IV.B.1, with an SSR value of 5.54×10^8 and an R^2 value of 0.2262. This finding is true for the models that handle the battle in one phase.

When the models which consider the change points are examined, it is seen that the model with the change point 7/7 is the one with the best fit, with an SSR value of 1.53×10^8 and an R^2 value of 0.7448.

Figure 78 shows the p and q values plotted for every model whose parameters are given in Table 42, except for the models with the change points since they have very large p and q parameters. The p and q values are also excluded for the model using the Morse-Kimball equations since these equations do not use p and q parameters.

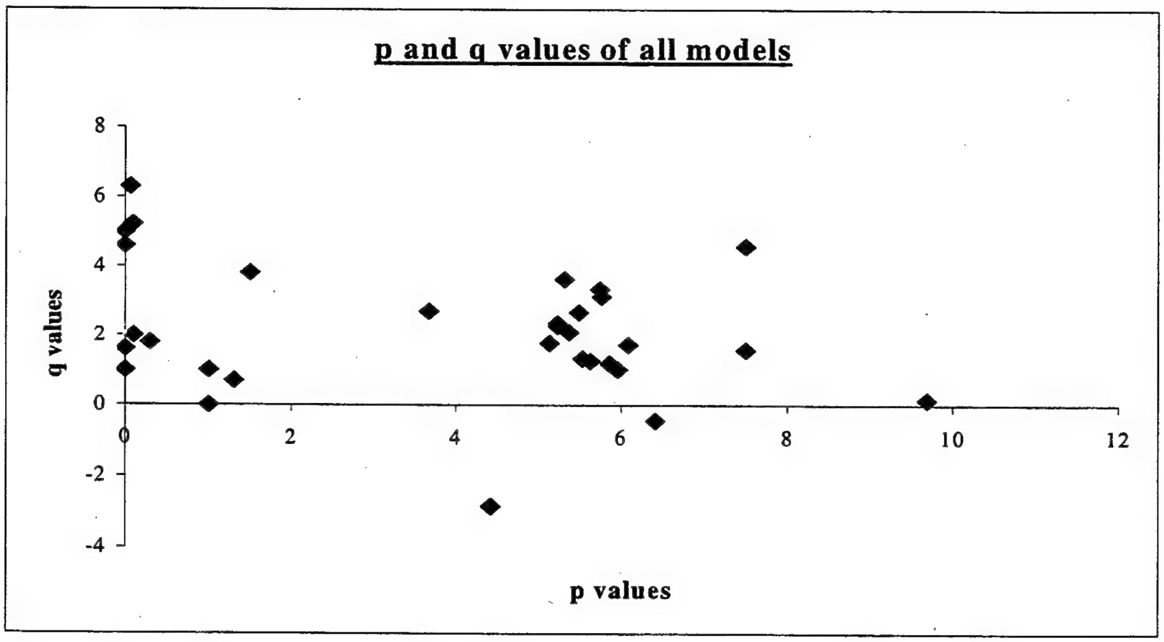

Figure 78. p and q parameters plotted for all the models given in Table 42.

When the pattern seen in Figure 78 is examined it is apparent that p and q parameters are clustered in two regions—one around the p=5-6, q=1-4 region, and the other around the q=1-6, p=0 region.

THIS PAGE IS LEFT INTENTIONALLY BLANK

V. CONCLUSIONS AND RECOMMENDATIONS

A. CONCLUSIONS

This thesis explores the applications of regression models on data from the World War II Battle of Kursk—the greatest single tank battle in history. The analysis tools used are two regression techniques, linear regression and robust LTS regression. The results and a brief interpretation of each regression model are given in the model's corresponding section. The results obtained from the statistical regression models are intended to provide insight into the Battle of Kursk, as well as into combat modeling in general.

When all the regression models are viewed together the following conclusions are reached.

- It is observed that the original Lanchester equations do not fit to the Battle of Kursk data, and therefore may not be appropriate for modeling the combat. Of the three ill-fitting Lanchester equations, the best fit is obtained by applying the linear law, which is used for modeling ancient warfare or area fire.
- The parameters derived from Bracken and Fricker's Ardennes studies do not apply to the Battle of Kursk data. This implies that there are no unique parameters that apply to all battles.
- Another interesting result with respect to Bracken's methodology and models is that, upon a closer examination of his findings, the SSR values given by multiple *p* and *q* values are in the same vicinity, as seen when plotted using a 3-D plot. This is clearly seen in Figures 79 through 82, which cover the breadth of approaches and show a variety of fits.

Figure 79 shows the SSR values plotted versus p and q parameters using Bracken's model, with the tactical parameter, for the Ardennes Campaign data when $a = 8x10^{-9}$, $b = 1x10^{-8}$ and d=1.25.

The SSR values do not change much in the vicinity of the best fit. Except for the spike seen on the upper far right corner, the SSR values are relatively insensitive to p and q values. Figure 80 gives a closer look at the region in which the lowest SSR value is found when p=1 and q=1. A broad range of parameters fit about the same in a valley of the surface as p increases and q decreases.

For Bracken's model, without the tactical parameter, for the Ardennes Campaign data when $a = 8x10^{-9}$, $b = 1x10^{-8}$, p=1.3 and q=0.7, the same pattern observed in Figures 79 and 80 also hold true, and this is true for other p and p analyses too.

Figures 81 and 82 show the 3-D grid plots for Bracken's model with the tactical parameter for the Battle of Kursk data when $a = 1.2x10^{-8}$, $b = 9x10^{-9}$ and d=0.9. Figure 82 provides a closer look at the region in which the lowest SSR value is found (i.e., when p=0.1 and q=2.0). This pattern is similar to what we have seen for Ardennes.

For Bracken's model, without the tactical parameter, for the Battle of Kursk data when $a = 1.2x10^{-8}$, $b = 9x10^{-9}$, p=0.3 and q=1.8, the same pattern observed in Figure 81 and Figure 82 also holds true. This result is again similar to what was found for the Ardennes data, suggesting that a broad range of possible models fit just about as well.

In the light of the findings stated above, it is a logical next step to have a look at the model's residual surface when a and b depend on p and q, and d=1. The results are given in Figures 81 through 84, where a and b parameters are chosen to minimize SSR

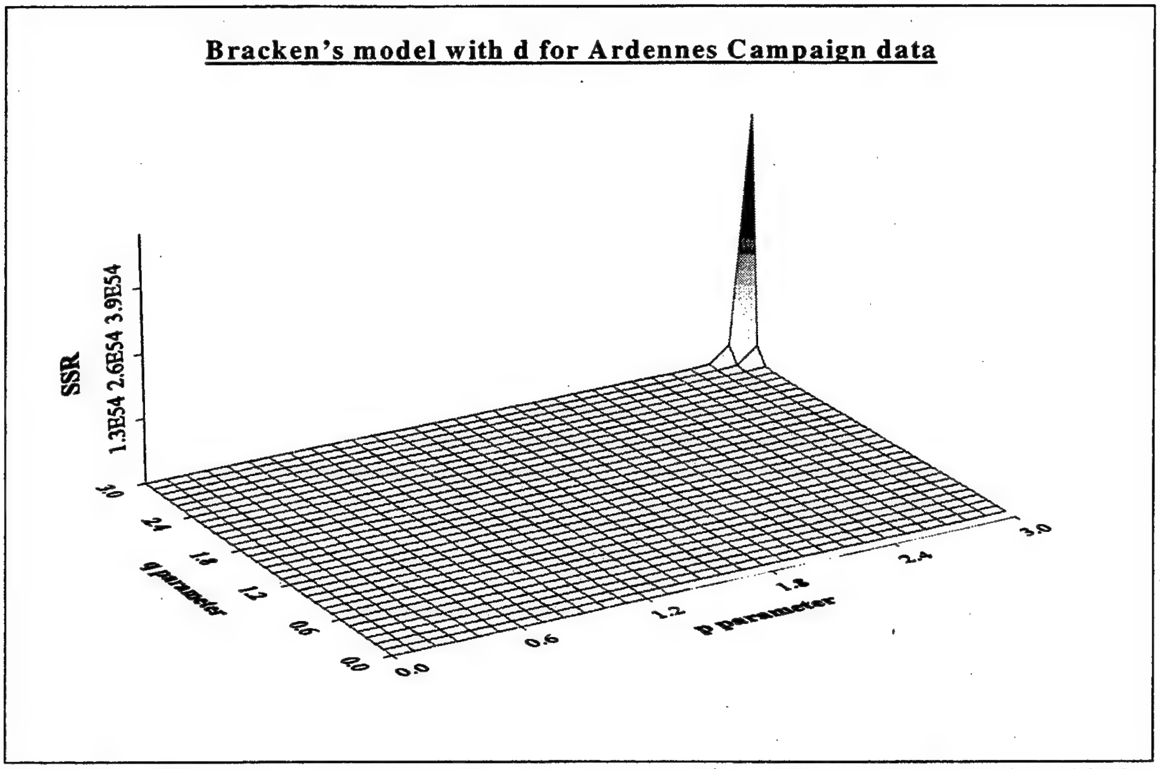

Figure 79. SSR values plotted versus p and q parameters using Bracken's model with the tactical parameter for Ardennes Campaign data.

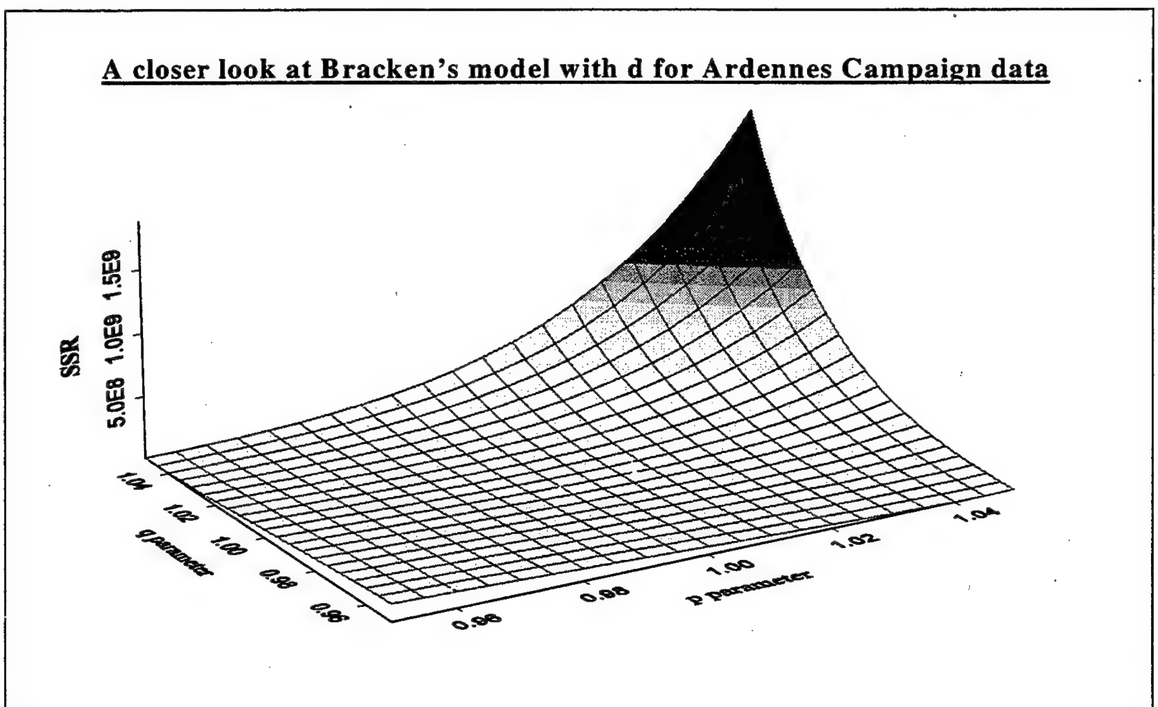

Figure 80. A closer look at the lowest SSR value for Bracken's model with the tactical parameter, for Ardennes Campaign data when p=1 and q=1.

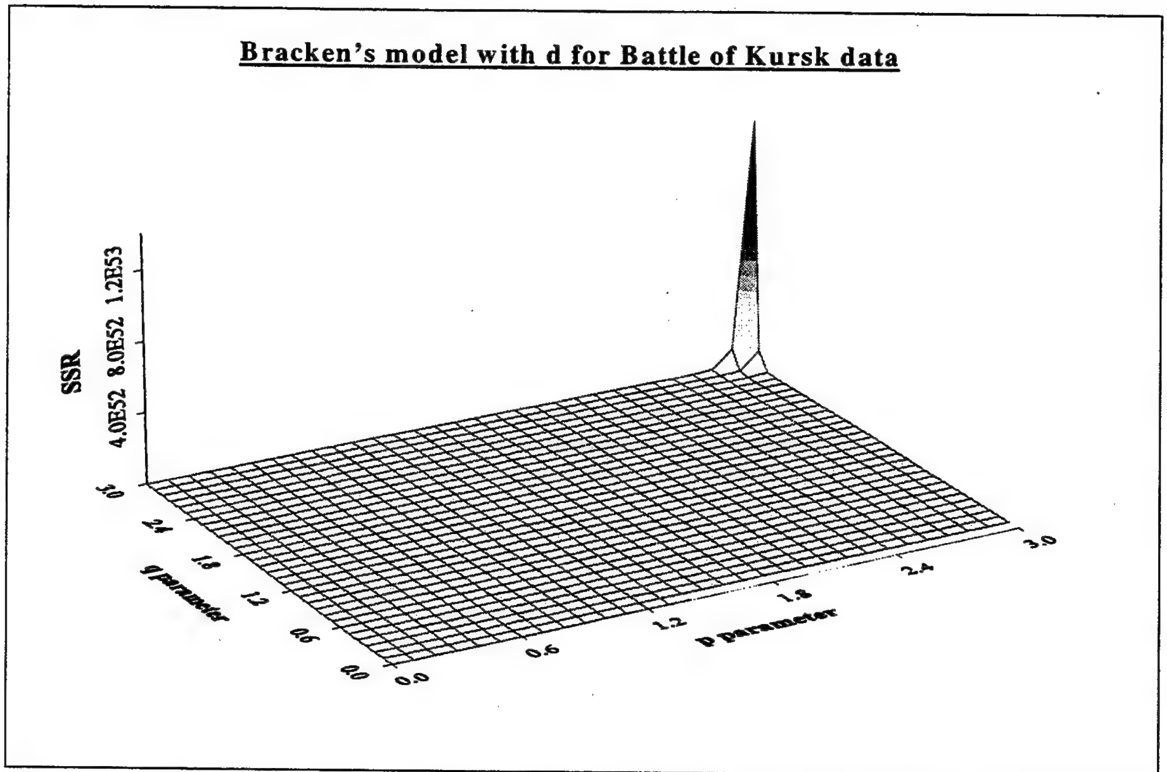

Figure 81. SSR values plotted versus p and q parameters using Bracken's model with the tactical parameter for Battle of Kursk data. This pattern holds true for other cases too.

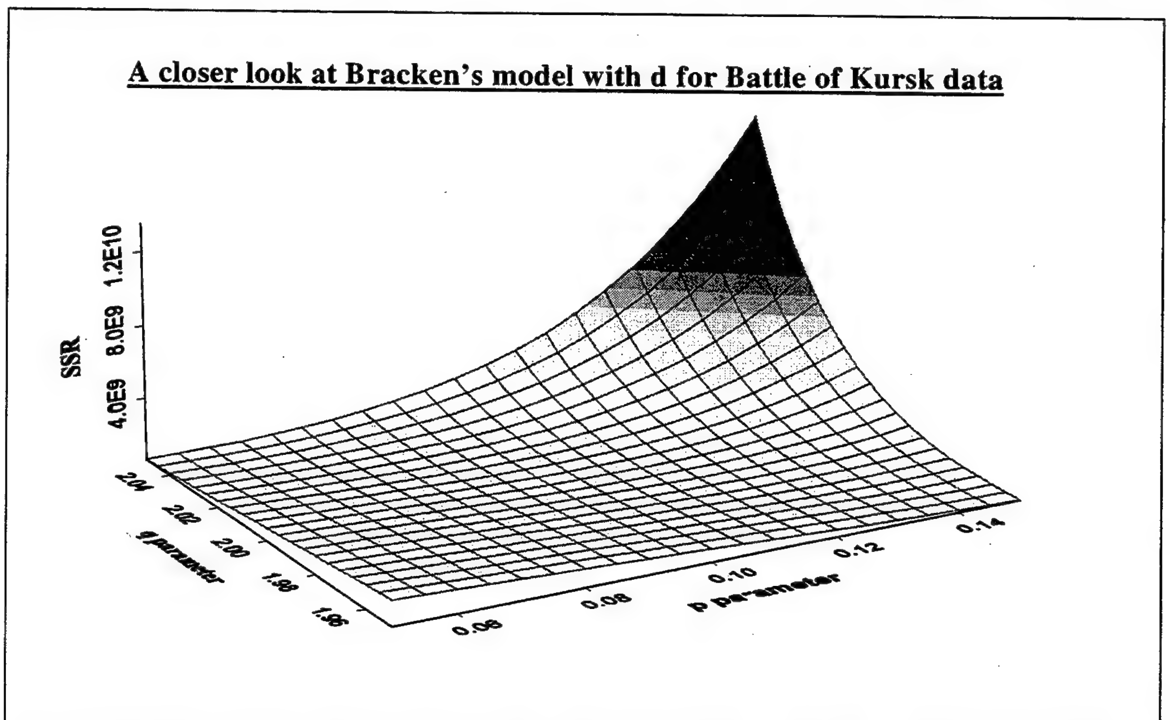

Figure 82. A closer look at the lowest SSR value for Bracken's model with the tactical parameter, for the Battle of Kursk data when p=0.1 and q=2.0. The observed pattern is similar to what is observed for Ardennes.

for regression through the origin given p and q. The best fitting a and b parameters are given as follows:

$$a = \frac{\sum_{i=1}^{n} \dot{B}(i)[R(i)^p B(i)^q]}{\sum_{i=1}^{n} [R(i)^p B(i)^q]^2} \tag{108}$$

$$b = \frac{\sum_{i=1}^{n} \dot{R}(i)[B(i)^p R(i)^q]}{\sum_{i=1}^{n} [B(i)^p R(i)^q]^2} \tag{109}$$

where i is the index of the days in a given battle, and n is the number of days in a given battle.

Figure 83 shows the 3-D plot of SSR values found for the Battle of Kursk data, where p values are varied between -0.5 and 10.0 with increments of 0.1, q values are varied between –1.0 and 3.0 with increments of 0.1, d=1.0, a and b values depend on p and q, and are determined by equations V.A.(108) and V.A.(109). Figure 85 shows the same area using a contour filled plot. A contour plot displays the contours of equally fitting p and q values in terms of SSR. This surface was generated with d fixed at 1.0. The models found in Fricker, Bracken and Clemens used a d parameter; hence their place on the surface does not necessarily measure the goodness of their fit. Furthermore, Fricker and Clemens used differently formatted data.

Figures 85 and 86 represent a detailed description of the region with the best fit. Figure 85 shows the 3-D plot of the SSR values found for the Battle of Kursk data, where p values are varied between 3.0 and 9.0 with increments of 0.1, q values are varied between 0.0 and 2.5 with increments of 0.1, d=1.0, a and b values depend on p and q, and

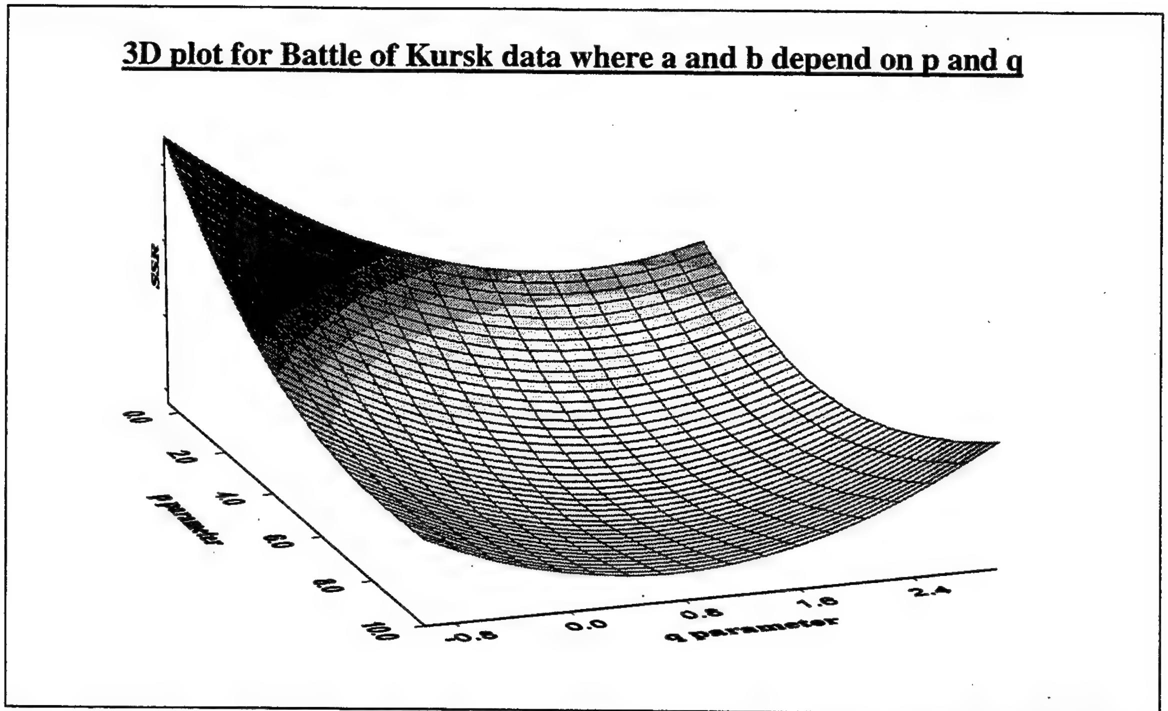

Figure 83. 3D plot of SSR values for Battle of Kursk data, p values are varied between –0.5 and 10.0 with increments of 0.1, q values are varied between –1.0 and 3.0 with increments of 0.1, d=1.0, a and b values depend on p and q.

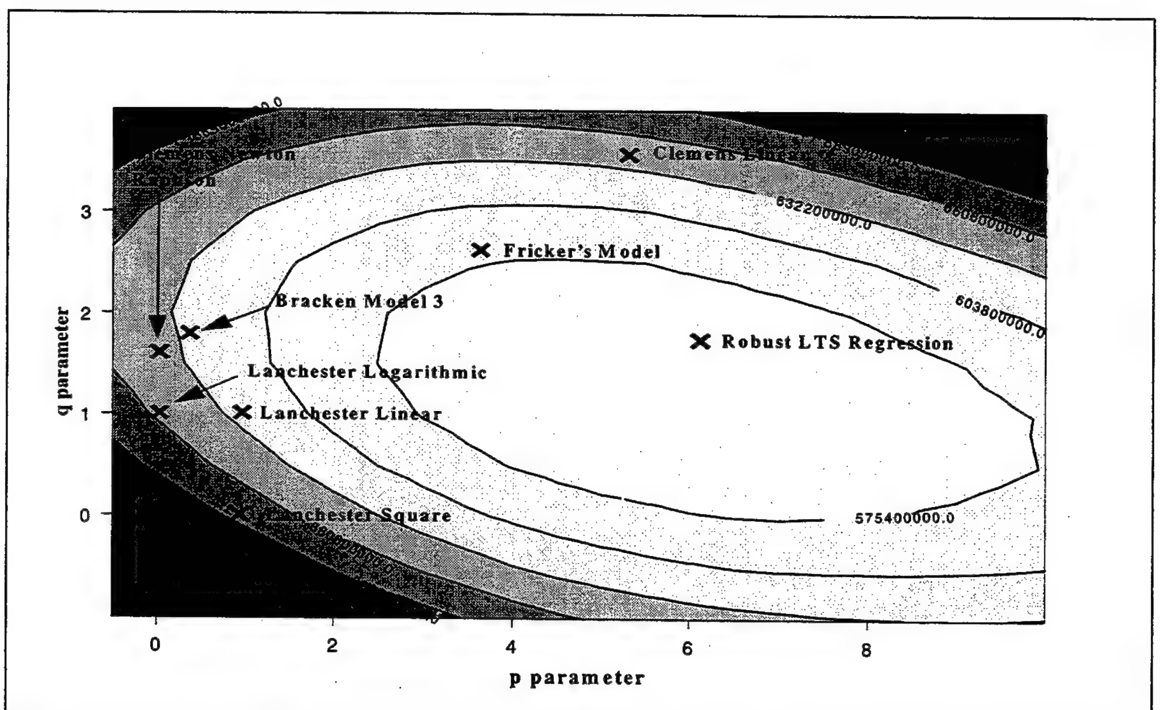

Figure 84. Contour filled plot of SSR values for Battle of Kursk data, p values are varied between –0.5 and 10.0 with increments of 0.1, q values are varied between –1.0 and 4.0 with increments of 0.1, d=1.0, a and b values depend on p and q. Also shown are each of the similar findings around the same area.

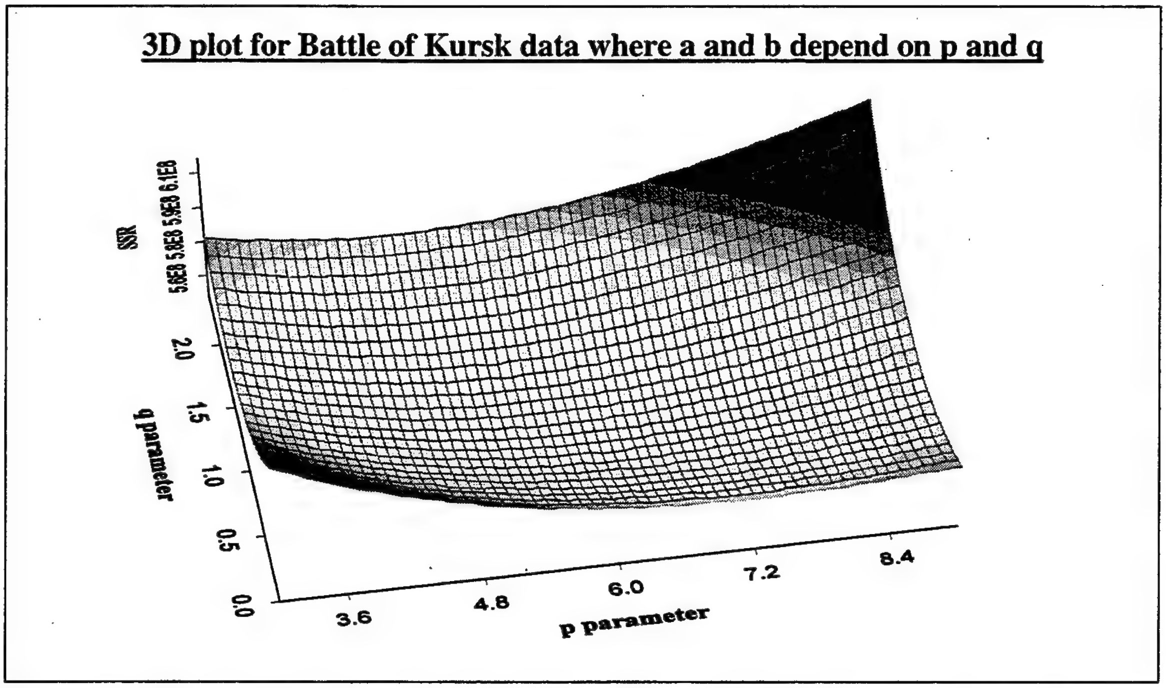

Figure 85. 3D plot of SSR values for Battle of Kursk data. p values are varied between 3.0 and 9.0 with increments of 0.1, q values are varied between 0.0 and 2.5 with increments of 0.1, d=1.0, a and b values depend on p and q.

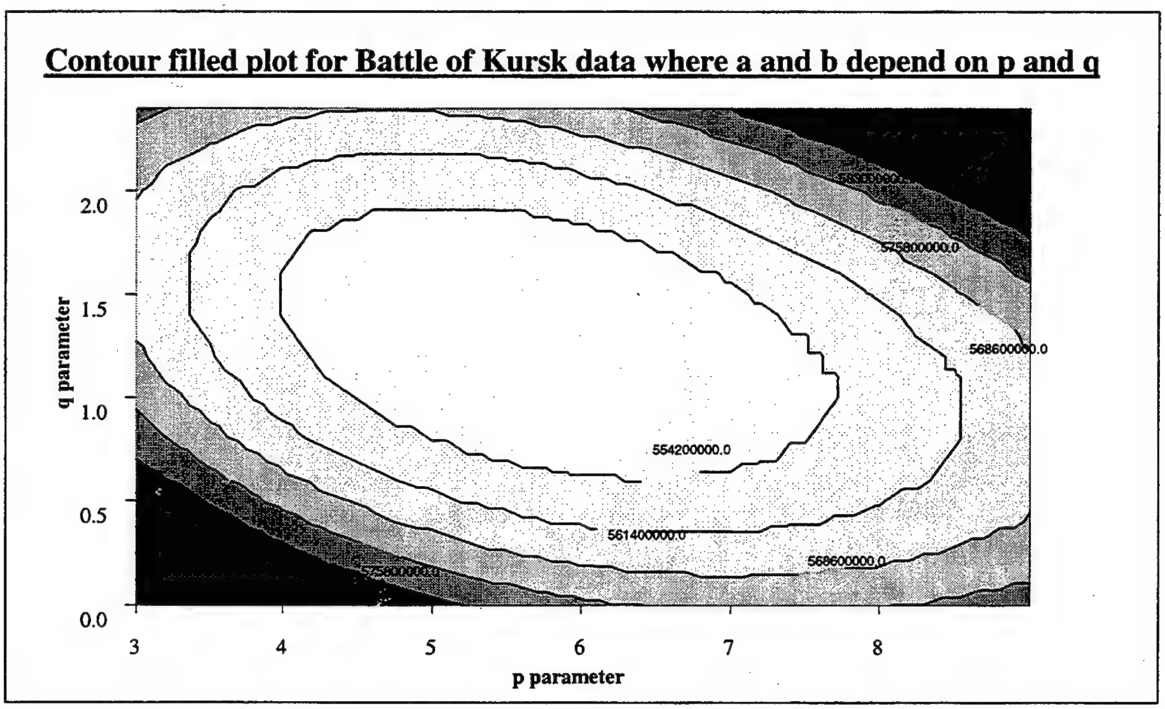

Figure 86. Contour filled plot of SSR values for Battle of Kursk data. p values are varied between 3.0 and 9.0 with increments of 0.1, q values are varied between 0.0 and 2.5 with increments of 0.1, d=1.0, a and b values depend on p and q.

are given as in equations V.A.(108) and V.A.(109). Figure 86 shows the same area using a contour filled plot.

The above results imply that there is no absolute best fit, as long as one stays in the broad vicinity of the identified best fit, it is likely to have similar fits, and there is not just one set of parameters that clearly gives a best fit. One can still find a similar fit as long as the estimated parameters are in the vicinity of the best fit. However, the area of the surface bounded by the lowest contour in Figure 86 says, roughly, the best fitting models have a q parameter between 0.5 and 2, while the p parameter is between 4 and 8. This observation is significant in that the Lanchester linear and square laws have a p value of 1 and the logarithmic law has a p value of 0. When p=1, the best fitting model has a 9% higher SSR value than the lowest found value of 5.54×10^{8}; which was found by using LTS regression.

A wide range of parameters fit equally well, but the question is, is this true for Ardennes campaign data too? Figure 87 shows the contour filled plot for the Ardennes data together with the best fits determined by Bracken and Fricker. Again, a and b depend on p and q, and d=1. The a and b parameters are chosen to minimize SSR for regression the through origin, and are given in equations V.A.(108) and V.A.(109). When Figure 87 is examined, one can see that the general pattern observed for the Kursk data is also observed for the Ardennes campaign data.

The bottom line conclusion is that different researchers using different methods all came up with very different answers because the surface around the models' fits is very flat.

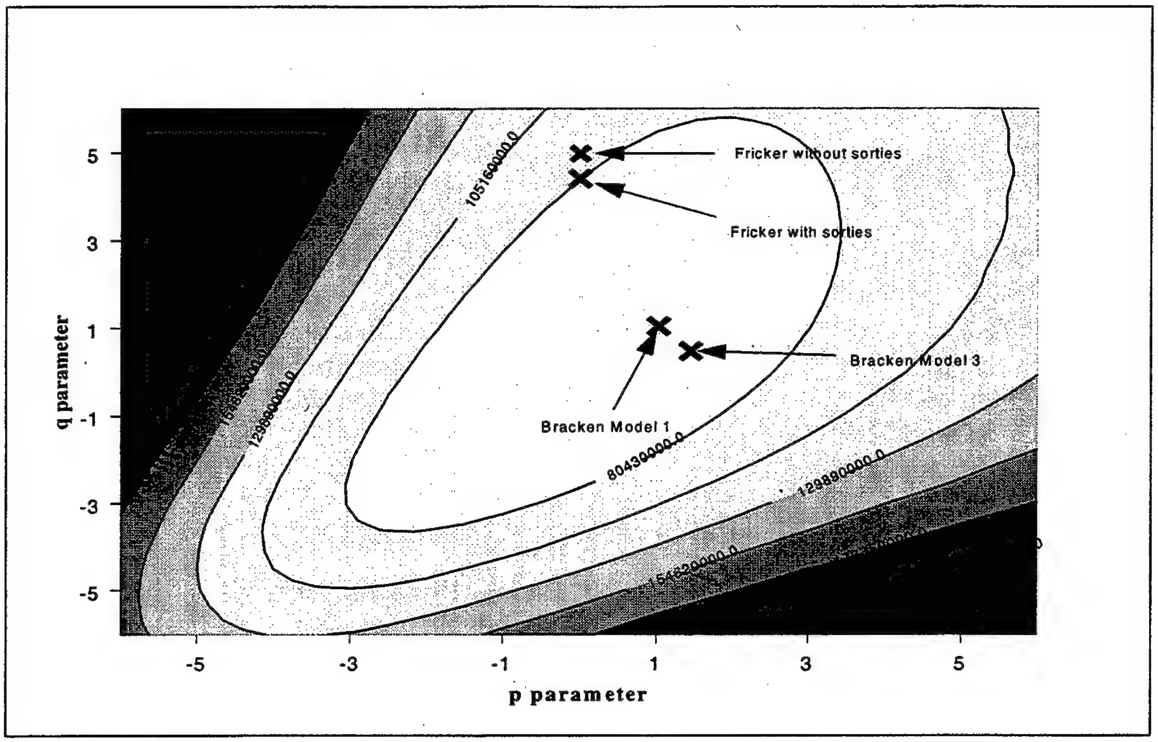

Figure 87. Contour filled plot of SSR values for Ardennes Campaign data. p values are varied between –6.0 and 6.0 with increments of 0.1, q values are varied between –6.0 and 6.0 with increments of 0.1, d=1.0, a and b values depend on p and q. Note: Fricker's method has a better R^2, this is not apparent from this figure which uses differently formatted data and no tactical parameter d.

- Fricker's methodology, when applied to Battle of Kursk data gives a better fit than Bracken's. This conclusion implies that the algorithm Fricker introduces in his study is useful in fitting the data and can be used in further studies.
- Throughout the study, with the exception of the two models whose results are given in equations IV.A.2.(14), IV.A.2.(15), IV.A.2.(18), IV.A.2.(19), and the models with the negative exponential parameters, the a parameter is always greater than the b parameter. This consistency implies that individually, German soldiers are more lethal than Soviet soldiers, and they also fought better than the Soviets. Also, according to this difference in parameters, German military expertise was much better than the Soviets' in the Battle of Kursk. This finding

was consistent throughout the battle. Despite their lack of military expertise, the Soviets won the battle due to their massive amount of supplies and manpower.

- Another significant result for the a and b parameters is that both are very small, and this result is consistent with Fricker's findings.
- The best fit to the data is observed when robust LTS regression model is applied. Robust LTS regression gives the smallest SSR value, which is 5.54×10^{8} when d=1.0. This finding was significant because it indicates no attacker/defender advantage.
- The d parameter, which gives the best fit using the linear regression model, is found to be 1.17. Using the a, b, p and q parameters found in equations IV.B.3.b.(40) and IV.B.3.b.(41), the data is analyzed in four distinct periods. The analysis revealed that it was usually advantageous to be the attacker in Battle of Kursk campaign. The only two days when the defender had the advantage was the first day when Germans attacked and the eighth day when the Soviets attacked. The Battle of Kursk was a major tank battle. Since a tank is an assault weapon, and is not optimally used as a defense weapon, the rationalization that the attacker will always have the advantage is considered a natural outcome of this battle.
- Finding only one tactical parameter d, and refusing to vary from that parameter through the battle is apparently a mistaken approach. The tactical parameter for a battle in which one side attacks a defender behind heavily fortified positions must not be the same with the tactical parameter for a battle in which one side is counterattacking and the other is making a hasty defense.

- The tactical parameter in the first half of the battle, being greater than the tactical parameter in the second half, indicates that the Germans were better than Soviets both in attack and defense. This finding is especially consistent with the tactical parameter value of 0.32 found on the eighth day, during which the Soviets counterattacked.
- The plots investigated in Section IV.B.7 show that if the force ratio is higher, then loss will be reduced because, as force ratio increases, loss decreases. This result is consistent with the force ratio approach, which is widely used in military simulation models today, showing the effectiveness and validity of the approach.
- The R^2 values given in Table 32 indicate that the model with the change point 7/7 represents the data with the best fit, with an R^2 value of 0.7748. The second best fit is observed with the model that divides the campaign in four different parts, with an R^2 value of 0.5689. This suggests that even an individual battle cannot be viewed as homogenous.
- Some models have negative R^2 values, meaning that one can have a better estimate of the attrition just by using the mean value, as opposed to using the model itself, and going into the modeling business. In other words, it is better to use the mean value for estimating the attrition instead of using the estimate given by the models, which have negative R^2 values. The negative R^2 values found in section IV.A.2 for Fricker's models occurred because the parameters were rounded off in Fricker [Ref.6]. Had more precise values been available the R^2 values would have been positive.

- Throughout the thesis, the robust LTS regression technique gives better fits to the data than the linear regression technique. This is because the robust regression models are useful for fitting linear relationships by discounting outlying data when the given data in hand contains significant outliers, as in our case.
- Combat models cannot provide clear-cut results to a military analyst. One cannot determine the outcome of a battle precisely by using combat models. Together with their use to gain insight about the battles and campaigns that happened in the past, combat models help to make better decisions by enabling the decision-maker to compare different alternatives using various combat modeling techniques.

B. RECOMMENDATIONS

The models presented in this thesis study do not include nor analyze total manpower data. Data for total manpower is present in the KDB and can be examined in the future studies.

The weights used for aggregating the forces are subject to research. A more complex model, one that includes the weights of weapons systems as unknown parameters to be estimated, can be set up and analyzed to find a better fit. And when the complex and numerous different weapon systems of today's military are considered, this shows potential to be a very interesting research topic.

Weapon systems other than the featured tanks in this study can be used to find a model with a better fit using a homogenous weapons scenario.

In this thesis, the change points for each side is on the same day. Another way to find a better fit would be to use different change points for each side, rather than using the same change point.

A final recommendation for the continued analysis of the Kursk database is to try to fit additional models other than the Lanchester models.

THIS PAGE IS LEFT INTENTIONALLY BLANK

APPENDIX A. DETAILED INFORMATION ON TYPES OF MANPOWER AND WEAPON SYSTEM LOSSES FOR BATTLE OF KURSK

A. TYPE OF MANPOWER CASUALTIES

Figure 88 shows the fraction of each type of casualty relative to total casualties. When all of four casualty types are considered, WIA accounted for the largest amount of casualties for both sides, and the German WIA fraction (0.751) was significantly higher than the Soviet WIA fraction (0.543). The next largest Soviet casualty fraction was for CMIA (0.230) while CMIA fraction accounted for the third largest German casualty fraction (0.031). The Soviet CMIA fraction was over 7 times greater than the German fraction. KIA accounted for the second largest German casualty fraction (0.15) while KIA fraction accounted for the third largest Soviet casualty fraction (0.217). Fewer than 1 percent of total casualties were DNBI for the Soviet (0.008), while DNBI accounted for almost 7 percent of total German casualties (0.65), which is over 7 times greater than the Soviet fraction.

Figure 89 shows the fraction of each type of casualty relative to initial OH Personnel. When all four casualty types are considered, WIA accounted for the largest amount of casualties for both sides again, and the Soviet WIA fraction (0.126) was slightly higher than the German WIA fraction (0.089). The next largest Soviet casualty fraction was for CMIA (0.053), while CMIA fraction accounted for the smallest German casualty fraction (0.003). The Soviet CMIA fraction was over 14 times greater than the German fraction (14.401). KIA accounted for the second largest German casualty fraction (0.018) while KIA fraction accounted for the third largest Soviet casualty fraction (0.050). While less than 1 percent of total casualties were DNBI for the both

sides, German DNBI fraction (0.007) was almost four (3.84) times higher than the Soviet DNBI fraction (0.002).

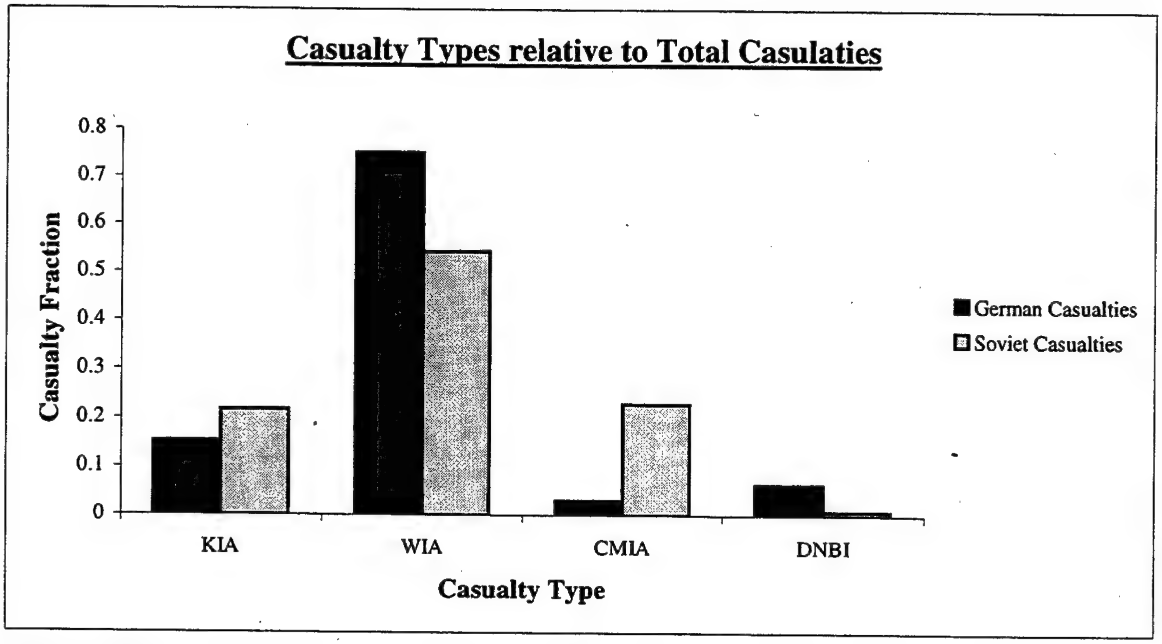

Figure 88. Fraction of personnel casualty types relative to total personnel casualties. WIA accounted for the largest amount of casualties for both sides.

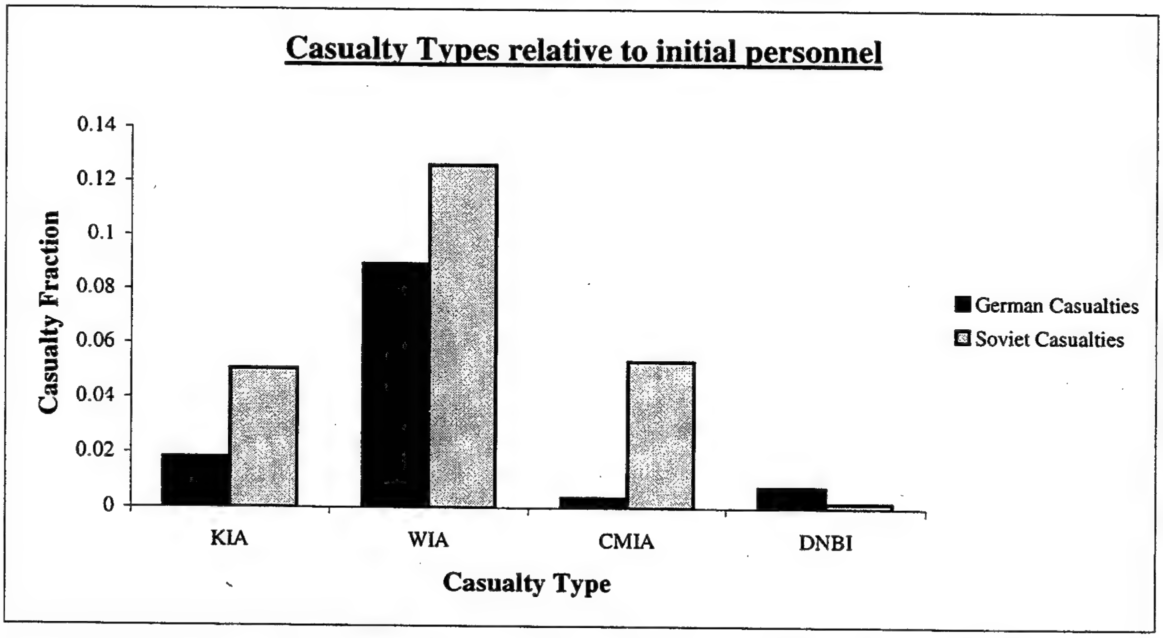

Figure 89. Fraction of personnel casualty types relative to total initial OH personnel. WIA accounted for the largest amount of casualties for both sides.

Figures 90 through 97 show daily and cumulative casualties for each casualty type of KIA, WIA, CMIA and DNBI consecutively. The largest differences are in KIA and CMIA. The Soviets had almost 5 (4.621) KIA for every one German KIA. The gap is even bigger for CMIA, with almost 24 (23.905) CMIA for every German CMIA. KIA and CMIA together, accounted for almost 45 (0.448) percent of total Soviet casualties, while they accounted for only slightly over 18(0.183) percent of total German casualties. For both sides, the majority of casualties were WIA. The Soviets had more than twice (2.343) as many WIA as the Germans.

The peak daily combat casualty rates occurred on July 5 and 12. The German peak daily rate was on July 5. The first day of the German attack was July 5, when only a minority of the Soviet force was engaged, and several Soviet units were overrun.

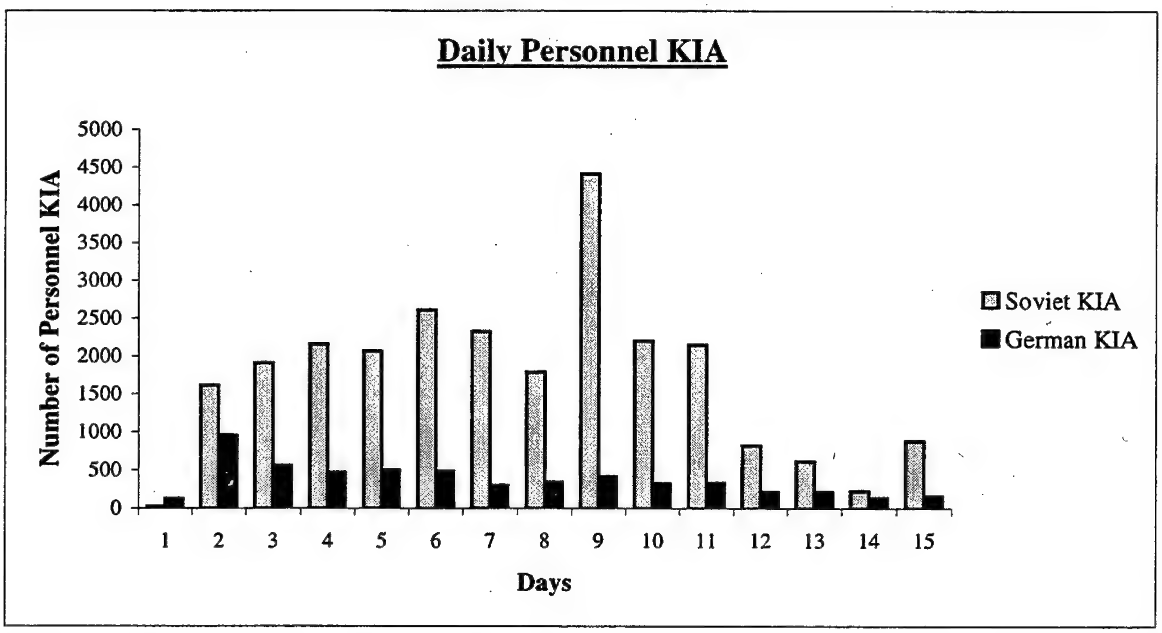

Figure 90. Daily number of total personnel casualties that are KIA. KIA denotes personnel that are killed in action.

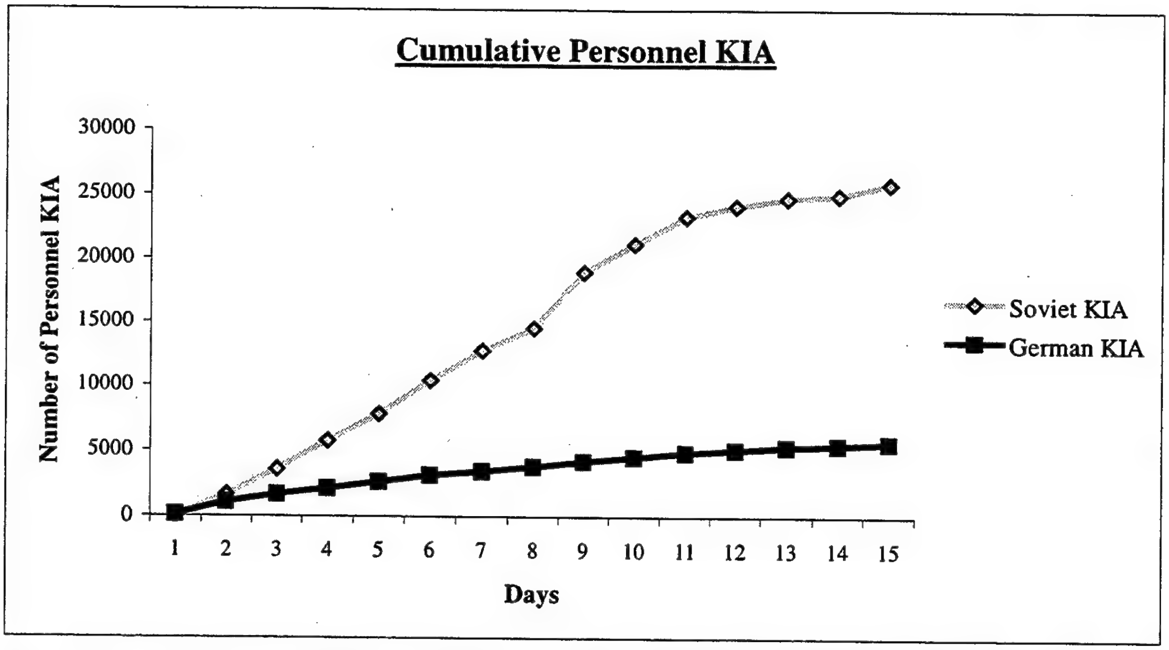

Figure 91. Daily cumulative number of total personnel casualties that are KIA. KIA denotes personnel killed in action.

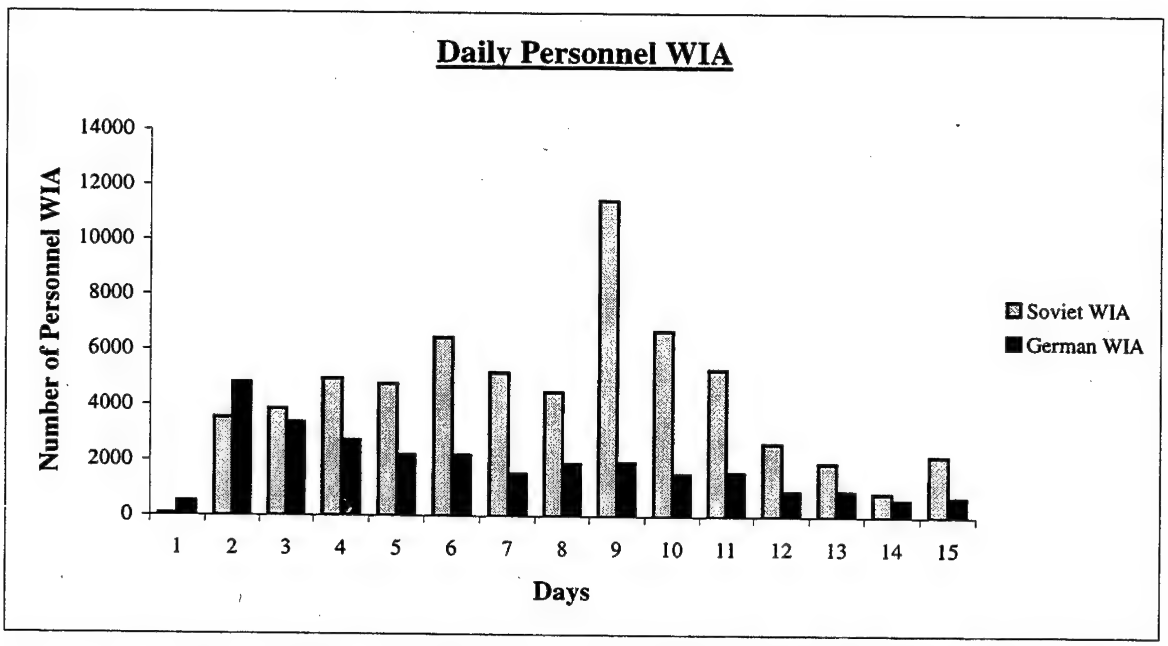

Figure 92. Daily number of total personnel casualties that are WIA. WIA denotes personnel wounded in action.

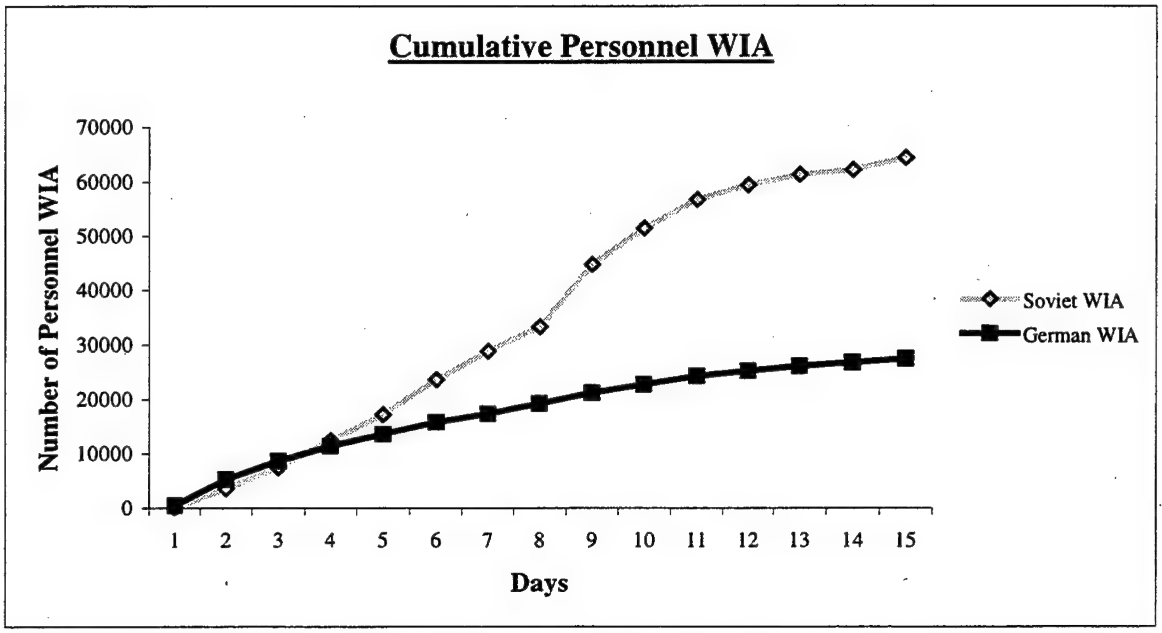

Figure 93. Daily cumulative number of total personnel casualties that are WIA. WIA denotes personnel wounded in action.

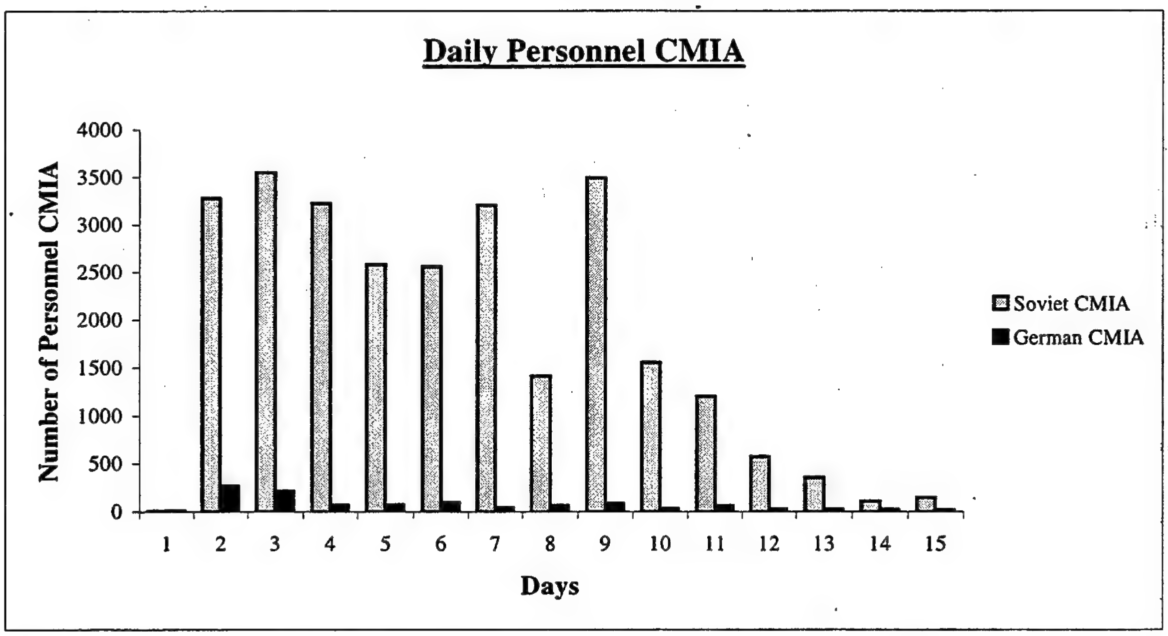

Figure 94. Daily number of total personnel casualties that are CMIA. CMIA denotes personnel captured or missing in action.

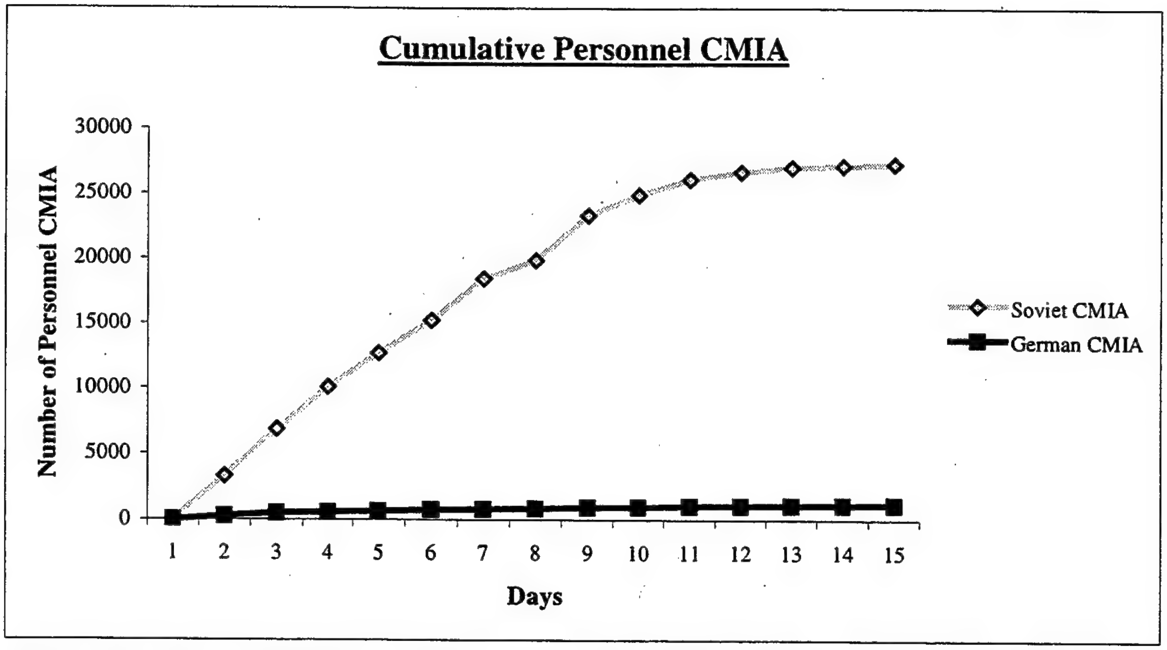

Figure 95. Daily cumulative number of total personnel casualties that are CMIA. CMIA denotes personnel captured or missing in action.

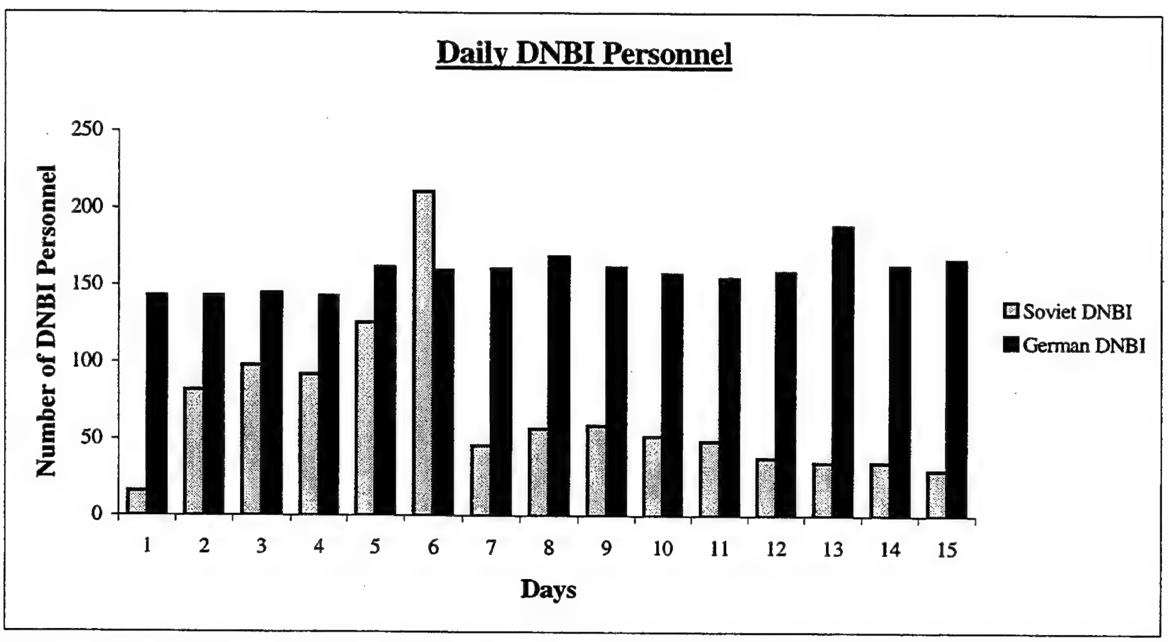

Figure 96. Daily number of total personnel casualties that are DNBI. DNBI denotes casualties due to disease and nonbattle injuries.

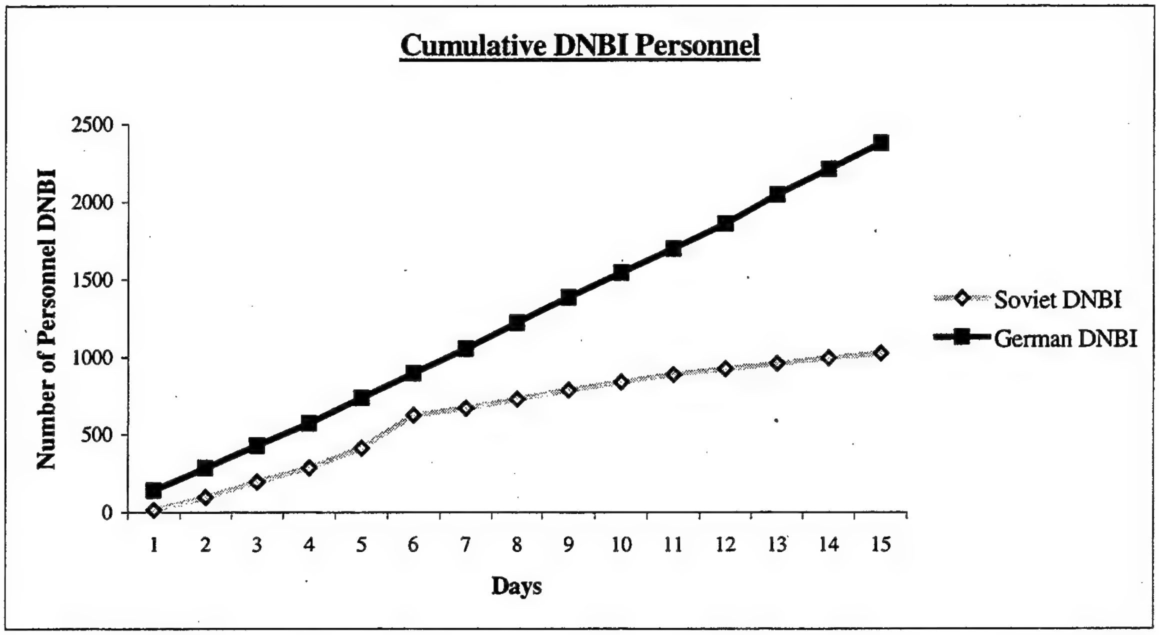

Figure 97. Daily cumulative number of total personnel casualties which are DNBI. DNBI denotes casualties due to disease and nonbattle injuries.

B. TYPE OF TANK LOSSES

Figure 98 shows the fraction of each type of tank loss relative to total tank losses. When both types of losses are considered, DAMAGED accounted for the largest amount of tank losses (0.849) for the German side, while DST+ABND accounted for the largest amount of tank losses (0.543) for the Soviet side. Consequently, DST+ABND accounted for the 15 (0.150) percent of tank losses for the Germans and DAMAGED accounted for the 54 (0.543) percent of tank losses for the Soviets. Overall, for every 1 DAMAGED Soviet tank, 1 (1.008) German tank was DAMAGED, and for every 1 DST+ABND German tank, almost 7 (6.655) Soviet tanks were DST+ABND.

Figure 99 shows the fraction of each type of tank loss relative to initial amount of OH tank. When both types of losses are considered, again DAMAGED accounted for the largest amount of tank losses for the German side, while DST+ABND accounted for the

largest amount of tank losses for the Soviet side. 89 (0.888) percent of the initial amount of OH German tank was DAMAGED, while only one sixth of that amount, i.e. 16 (0.157) percent, was DST+ABND. Fifty (0.495) percent of the initial amount of OH Soviet tanks was DST+ABND, while 42 (0.415) percent, was DST+ABND.

Figures 100 through 103 show daily and cumulative tank losses for each type of tank losses, namely DST+ABND and DAMAGED consecutively.

C. TYPE OF APC LOSSES

Figure 104 shows the fraction of each type of APC loss relative to total APC losses. When both types of losses are considered, DAMAGED accounted for the largest amount of APC losses (0.739) for the Germans, while DST+ABND accounted for the

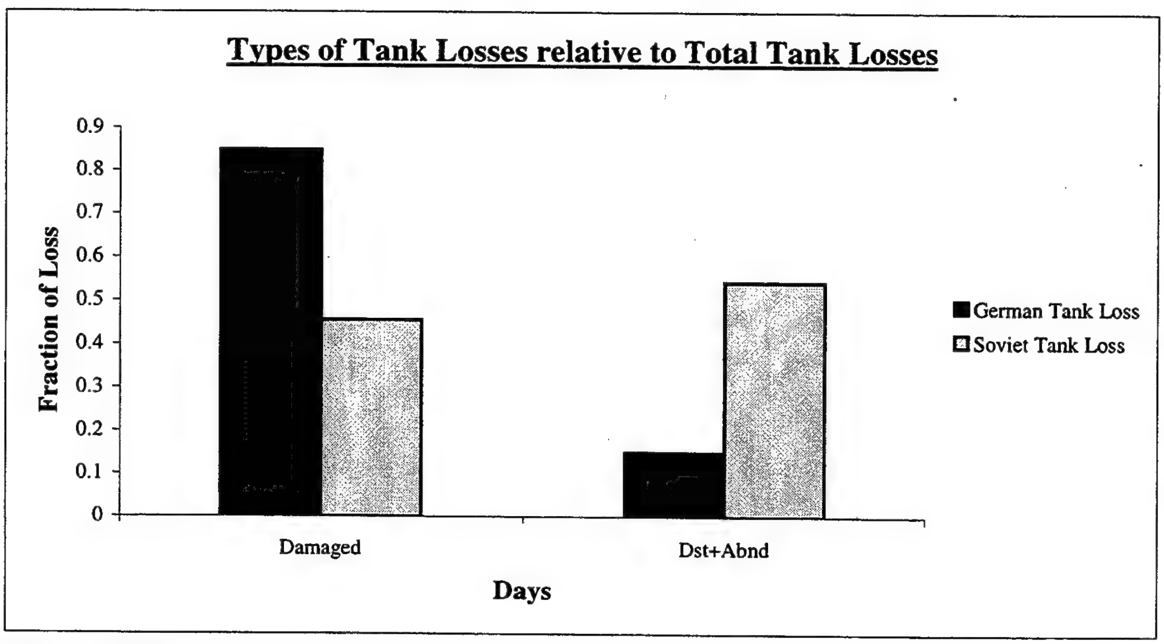

Figure 98. Fraction of each type of tank loss relative to total tank losses. When both types of losses are considered, DAMAGED accounted for the largest amount of tank losses for the German side, while DST+ABND accounted for the largest amount of tank losses for the Soviet side

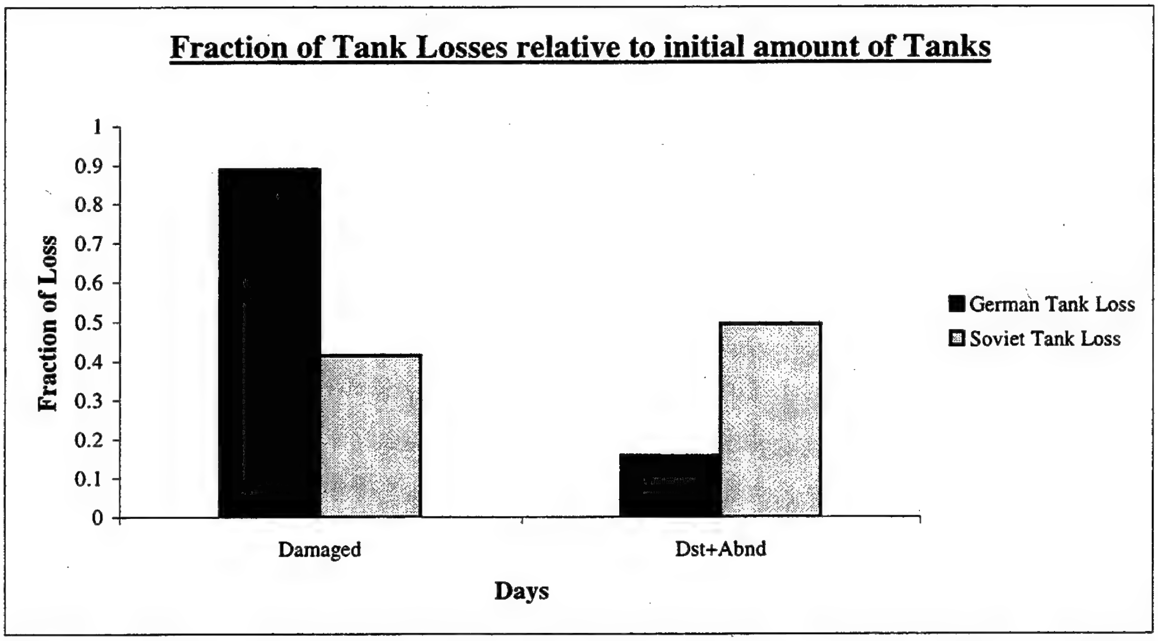

Figure 99. Fraction of each type of tank loss relative to initial number of OH tanks. When both types of losses are considered, DAMAGED accounted for the largest amount of tank losses for the German side, while DST+ABND accounted for the largest amount of tank losses for the Soviet side.

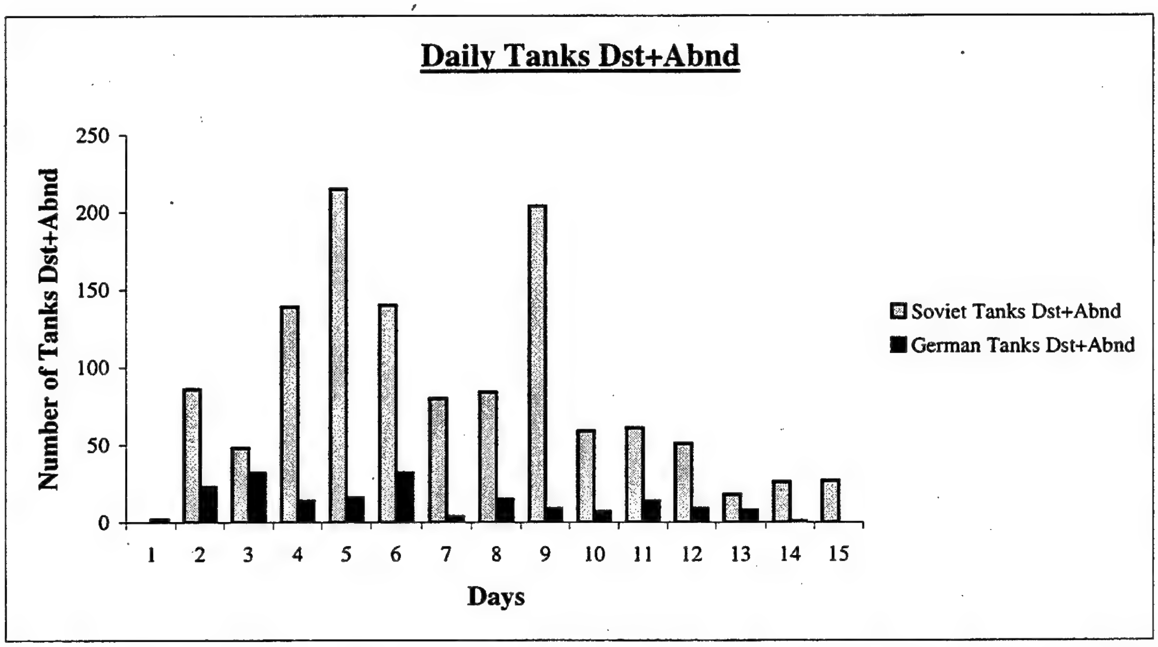

Figure 100. Daily number of total tank losses that are DST+ABND. DST+ABND denotes the weapons that are destroyed or abandoned. Soviets had no tanks that are DST+ABND on day 1, Germans had no tanks which are DST+ABND on day 15.

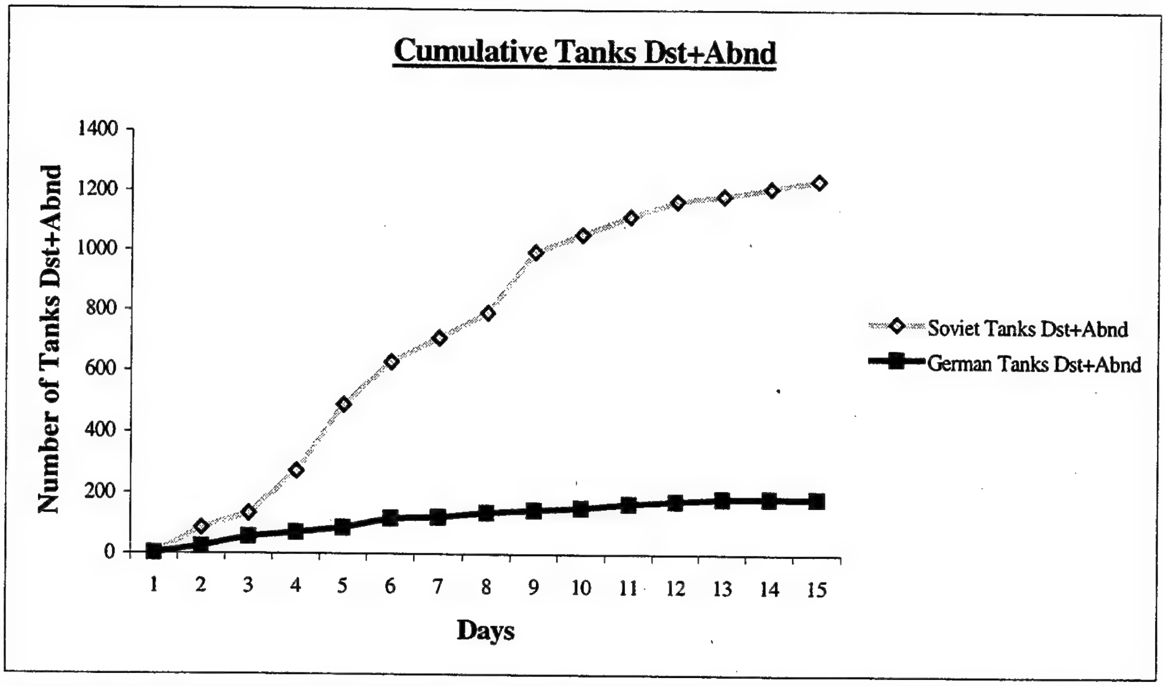

Figure 101. Cumulative number of tank losses that are DST+ABND. DST+ABND denotes the weapons that are destroyed and abandoned.

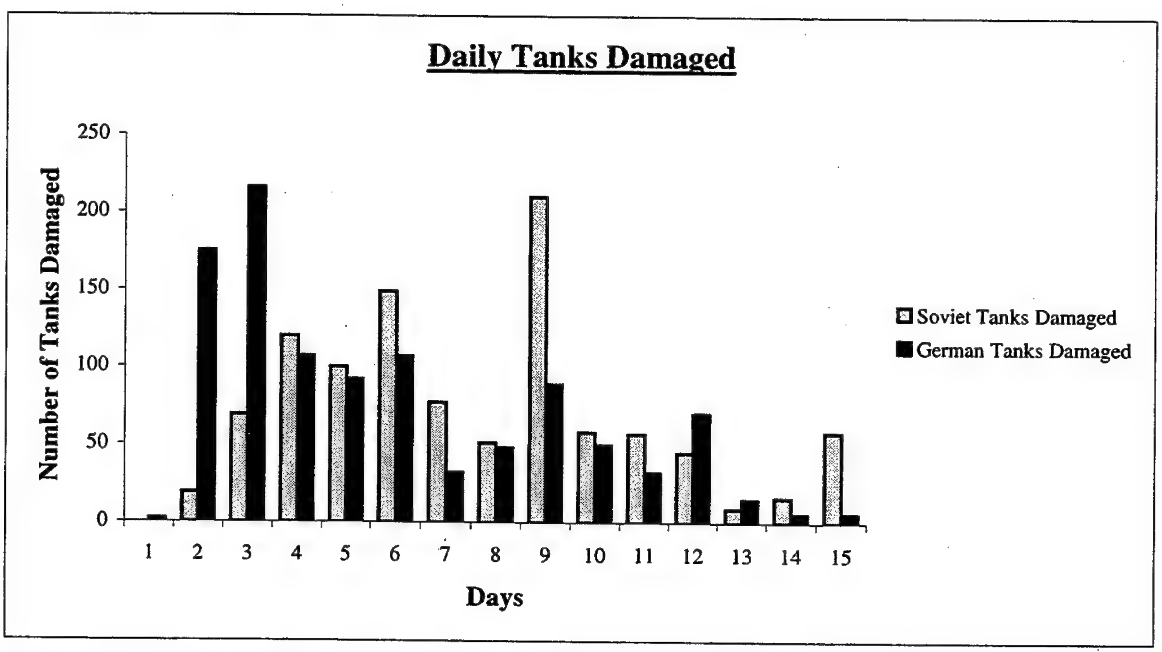

Figure 102. Daily number of total tank losses that are damaged. Soviets had no damaged tanks on day 1.

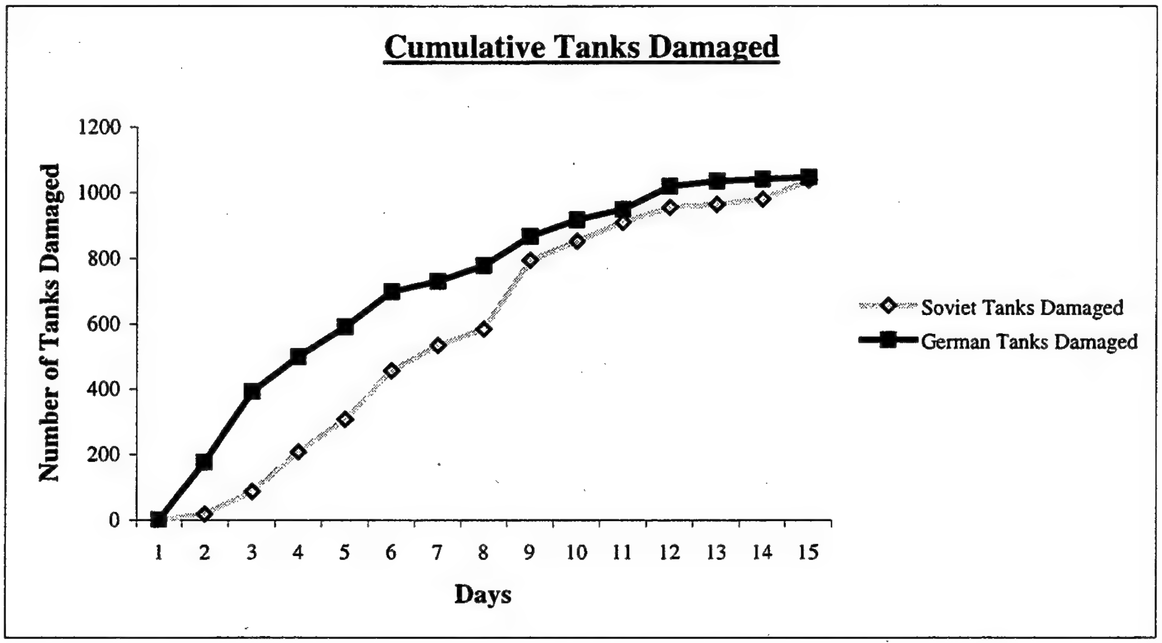

Figure 103. Daily cumulative number of total tank losses that are damaged.

largest amount of APC losses (0.623) for the Soviet side. Consequently, DST+ABND accounted for the 26 (0.260) percent of APC losses for the Germans and DAMAGED accounted for the 38 (0.376) percent of APC losses for the Soviets. Overall, for every 1 DAMAGED Soviet APC, more than 3 (3.227) German APCs were DAMAGED, and for every 1 DST+ABND German APC, 1.46 Soviet APCs were DST+ABND.

Figure 105 shows the fraction of each type of APC loss relative to initial amount of OH APC. When both types of losses are considered, again DAMAGED accounted for the largest amount of APC losses for the German side, while DST+ABND accounted for the largest amount of APC losses for the Soviet side. Twelve (0.121) percent of the initial amount of OH German APC were DAMAGED, while only one third of that amount, i.e. 4 (0.042) percent, were DST+ABND. Fourteen (0.142) percent of the initial

amount of OH Soviet APC were DST+ABND, while 9 (0.086) percent, were DST+ABND.

Figures 106 through 109 show daily and cumulative APC losses for each type of APC losses, namely DST+ABND and DAMAGED consecutively.

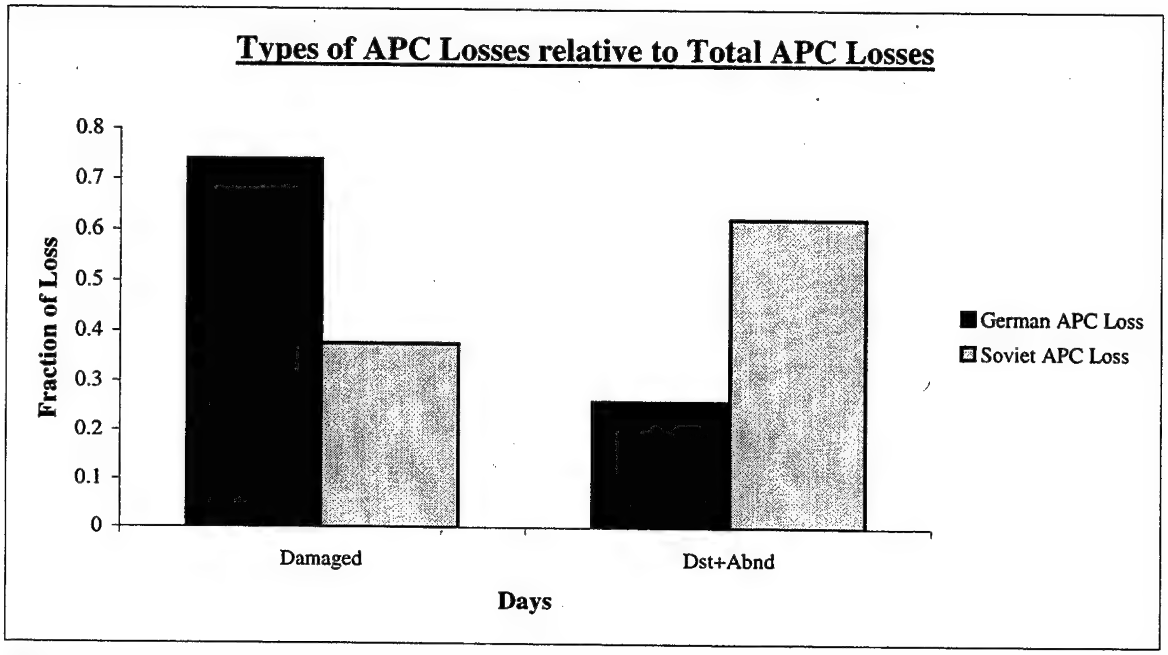

Figure 104. Fraction of each type of APC loss relative to the total APC losses. DAMAGED accounted for the largest amount of APC losses for the German side, while DST+ABND accounted for the largest amount of APC losses for the Soviet side.

D. TYPE OF ARTILLERY LOSSES

Figure 110 shows the fraction of each type of Artillery loss relative to total artillery losses. When both types of losses are considered, DST+ABND accounted for the largest amount of artillery losses for both sides, and the Soviet DST+ABND fraction (0.847) was significantly higher than the German fraction (0.559). Consequently, DAMAGED accounted for the 44 (0.440) percent of artillery losses for the Germans and 15 (0.152) percent of artillery losses for the Soviets. Overall, for every 1 DAMAGED Soviet Artillery, nearly 3 (2.545) German artillery was DAMAGED, and for

every 1 DST+ABND German artillery, almost 2 (1.718) Soviet artillery was DST+ABND.

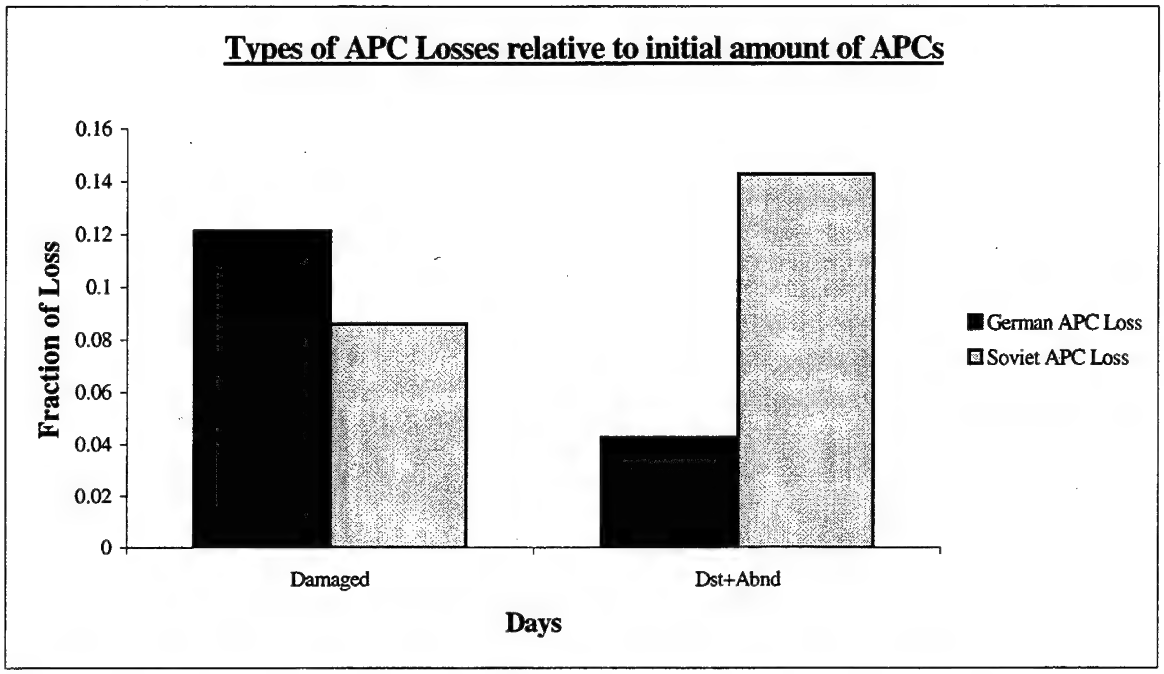

Figure 105. Fraction of each type of APC loss relative to the initial amount of APCs.

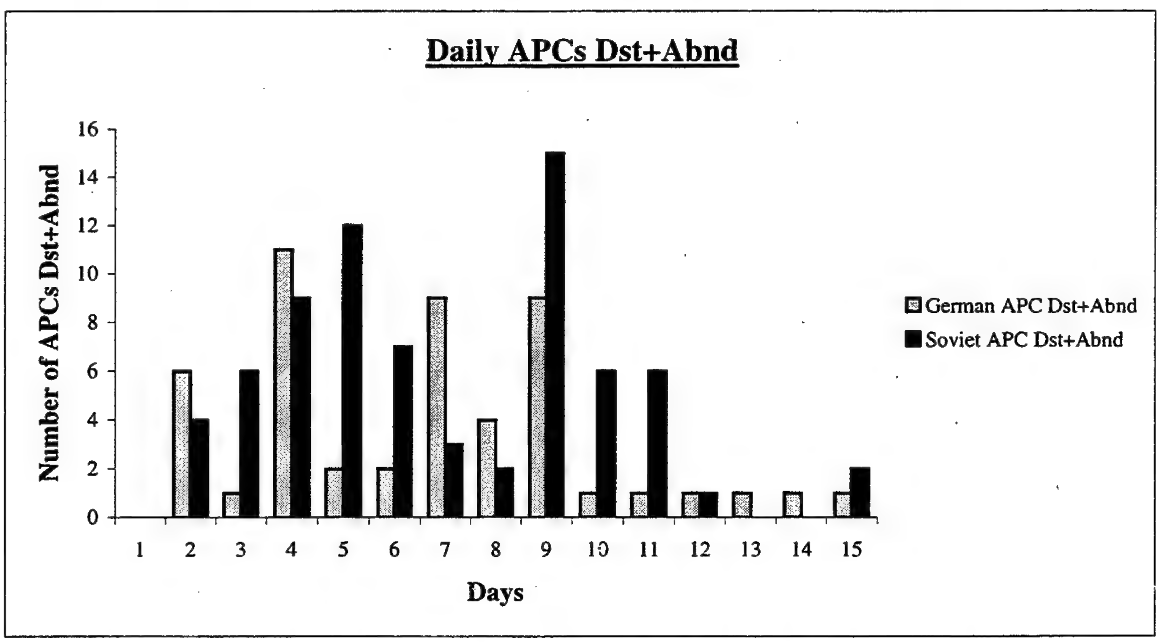

Figure 106. Daily number of total APC losses that are DST+ABND. DST+ABND denotes weapons that are destroyed or abandoned. Soviets had no APCs that are DST+ABND on days 1,13,14. Germans had no APCs that are DST+ABND on day 1.

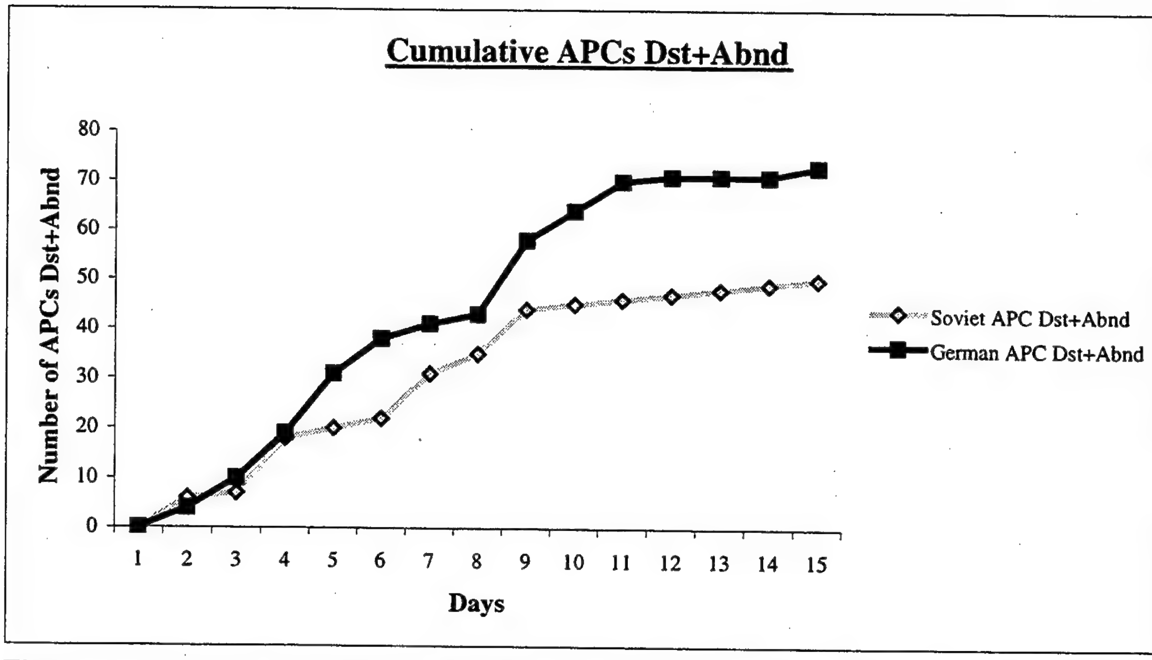

Figure 107. Daily Cumulative number of total APC losses that are DST+ABND. DST+ABND denotes weapons that are destroyed or abandoned.

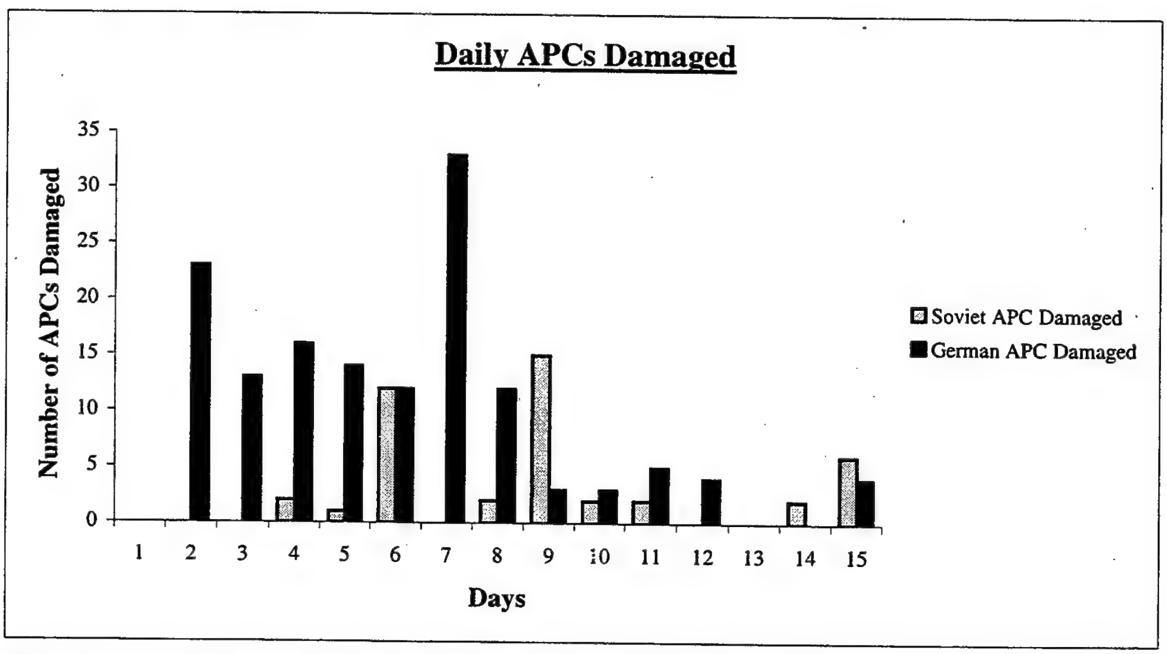

Figure 108. Daily number of Total APC losses that are damaged. Soviets had no damaged APCs on days 1, 2, 3, 7, 12 and 13. Germans had no damaged APCs on days 1, 13 and 14.

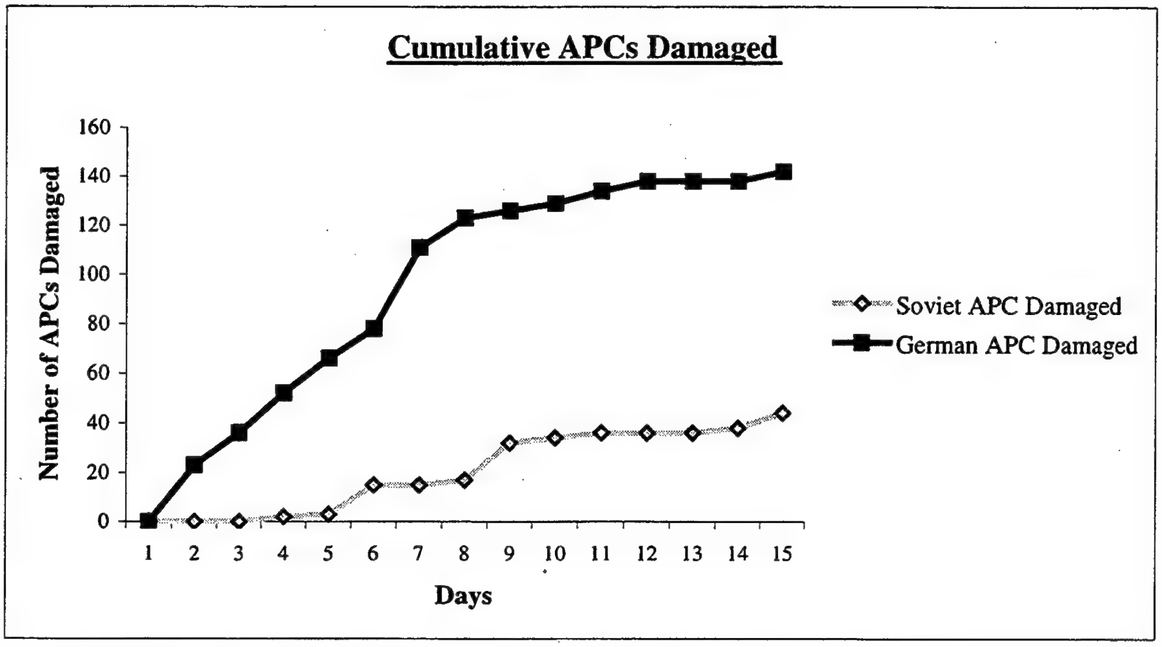

Figure 109. Daily cumulative number of total APC losses that are damaged.

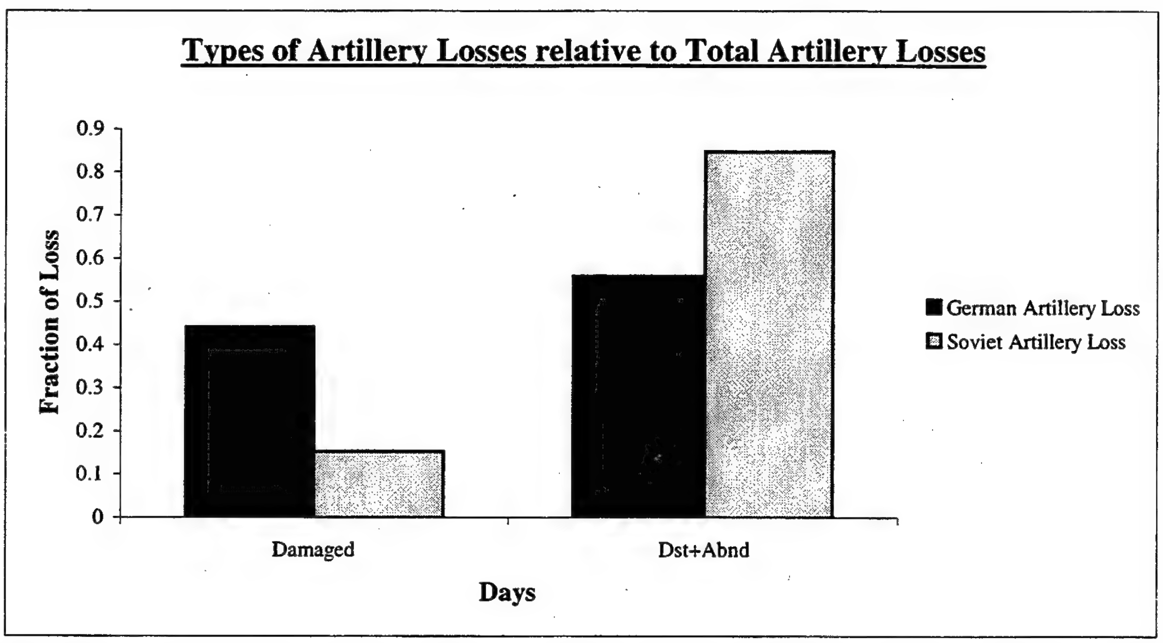

Figure 110. Fraction of each type of artillery losses relative to total artillery losses. DST+ABND accounted for the largest amount of artillery losses for both sides.

Figure 111 shows the fraction of each type of artillery loss relative to initial amount of OH artillery. When both types of losses are considered, again DST+ABND accounted for the largest amount of artillery losses both for the German side and also for the Soviet side. 6 (0.059) percent of the initial amount of OH German artillery was DST+ABND, while 5 (0.047) percent was DAMAGED. 17(0.169) percent of the initial amount of OH Soviet artillery was DST+ABND, while only almost one sixth of that amount, i.e. 3 (0.030) percent, was DAMAGED.

Figures 112 through 115 show daily and cumulative Artillery losses for each type of artillery losses namely DST+ABND and DAMAGED consecutively.

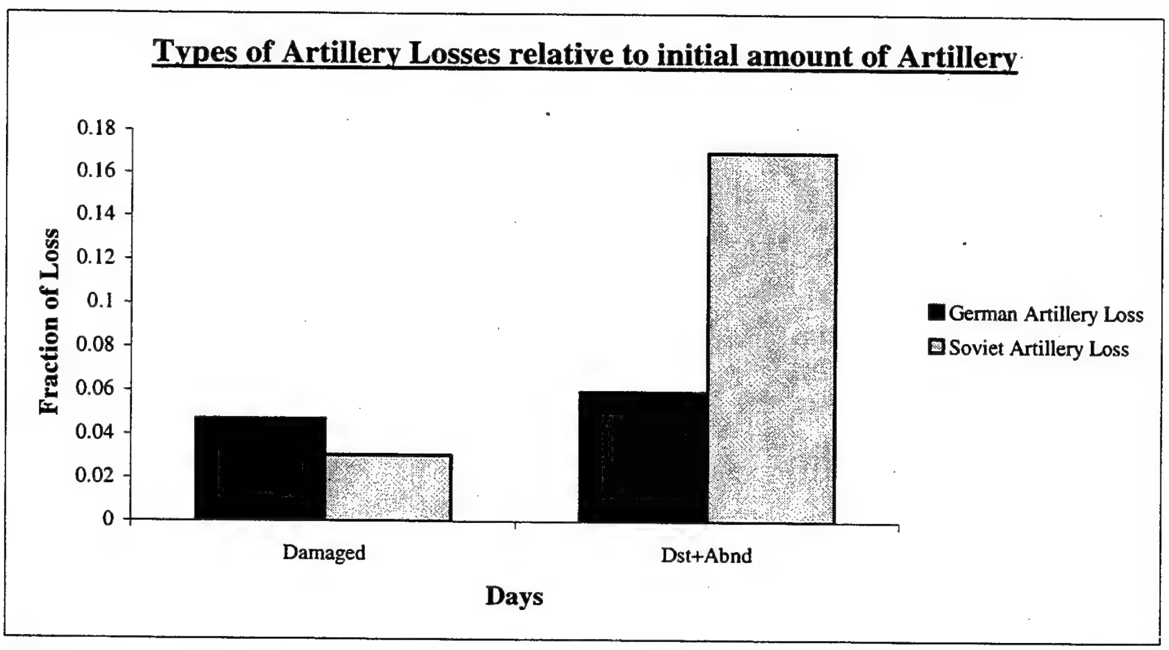

Figure 111. Fraction of each type of loss relative to initial amount of artillery. DST+ABND accounted for the largest amount of artillery losses for both the German side and the Soviet side.

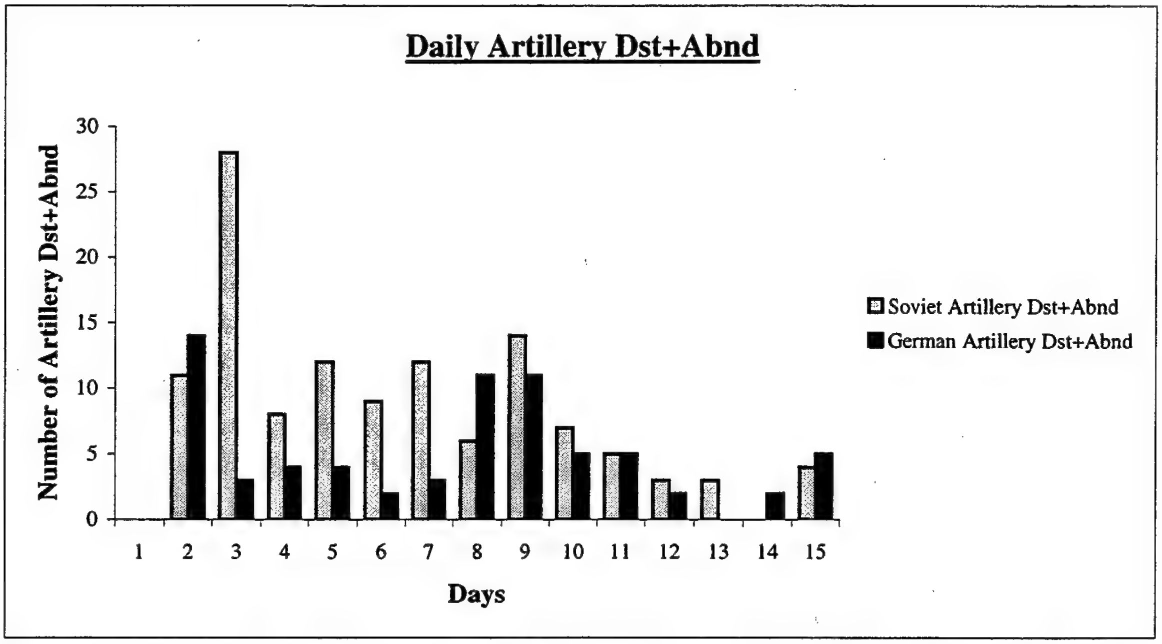

Figure 112. Daily number of total artillery losses which are DST+ABND. DST+ABND denotes weapon systems that are destroyed or abandoned.

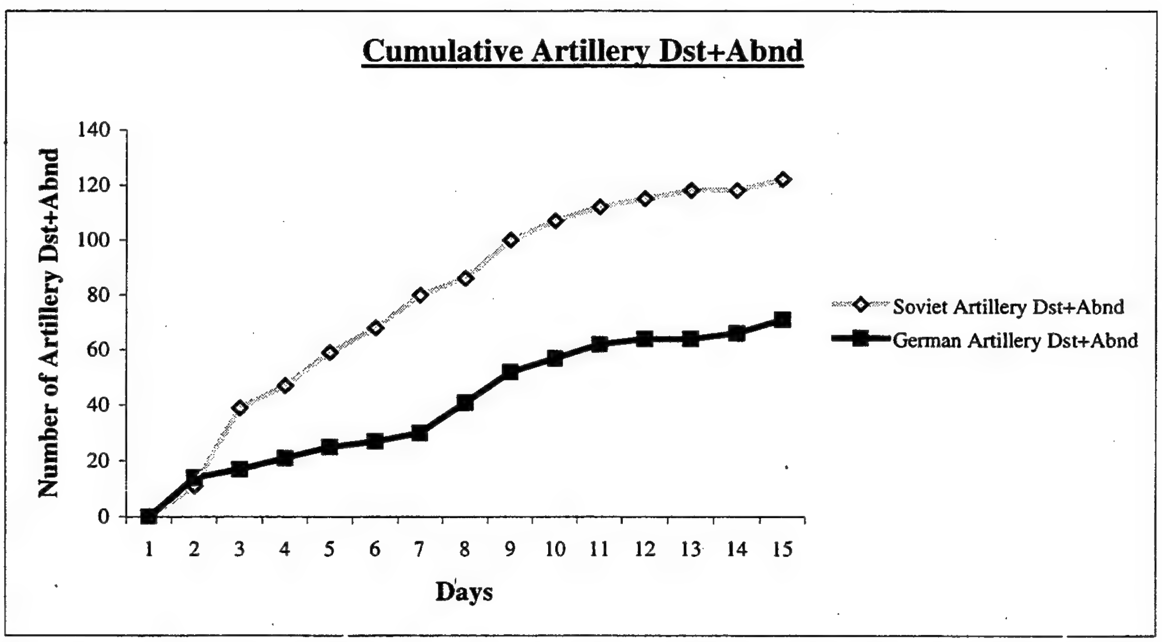

Figure 113. Daily cumulative number of total artillery losses that are DST+ABND. DST+ABND denotes weapon systems that are destroyed or abandoned.

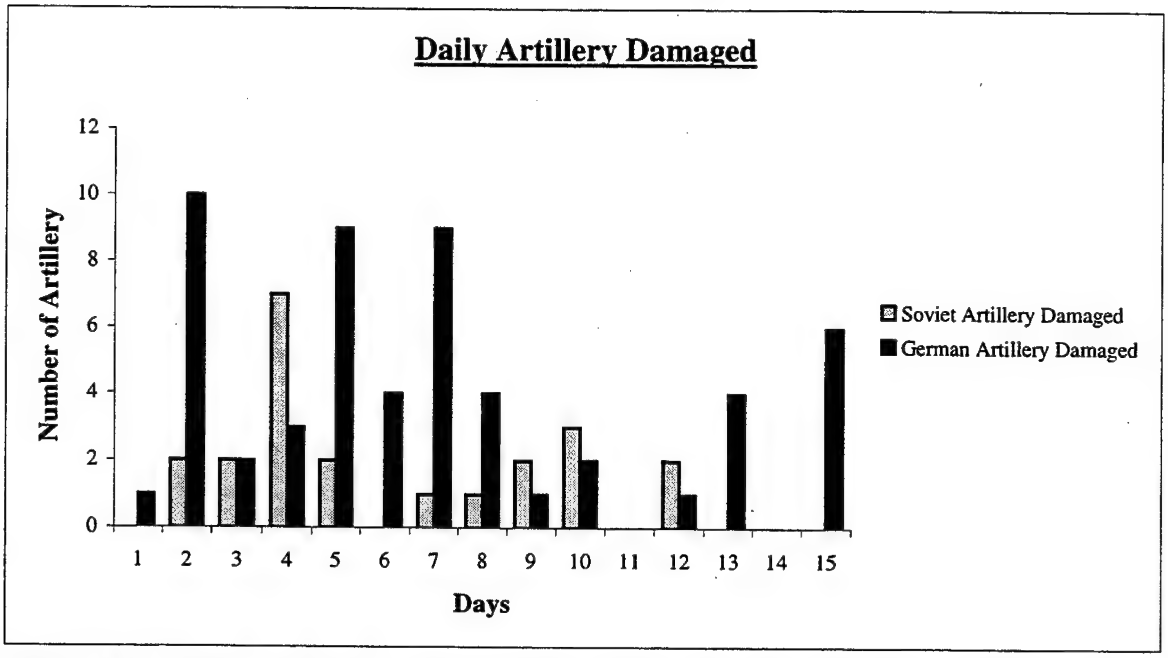

Figure 114. Daily number of total artillery losses that are damaged. Soviets had no damaged artillery on days 1, 6, 11, 13, 14 and 15. Germans had no damaged artillery on days 11 and 14.

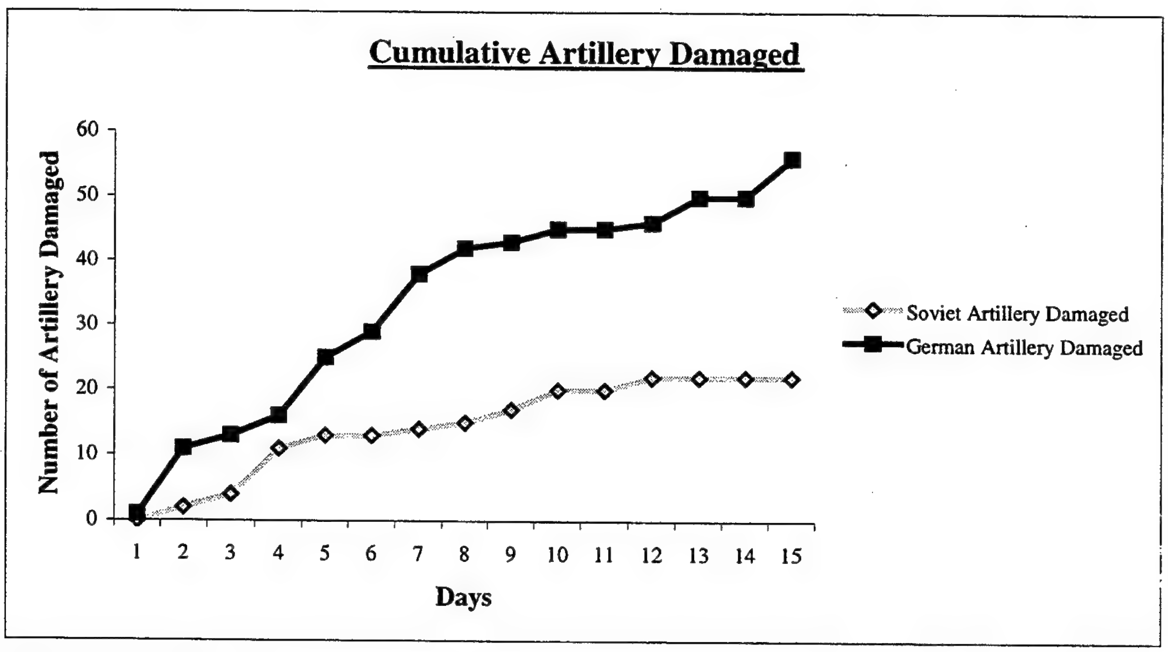

Figure 115. Daily cumulative number of total artillery losses that are damaged.

LIST OF REFERENCES

1. Merriam-Webster Dictionary "definition of war." [http://www.britannica.com/cgi-bin/dictionary?va=war]. January 2000.

2. Osipov, M., *The influence of the Numerical Strength of Engaged Sides on their Casualties*, Voenniy Sbornik (Military Collection), 1915.

3. Merriam-Webster Dictionary "Lanchester, Frederick William." [http://www.britannica.com/bcom/eb/article/5/0,5716,48115+1,00.html]. January 2000.

4. Lanchester, F. W., *Aircraft in Warfare: The Dawn of the Fourth Arm*, Constable, London, 1916.

5. Hughes, W. P., *Military Modeling*, 2nd Edition, p.77, MORS, 1989.

6. Fricker, R. D., *Attrition Models of the Ardennes Campaign*, Naval Research Logistics, vol. 45, no. 1, pp. 1-22, 1998.

7. Taylor, J. G., *Lanchester Models of Warfare* (two volumes), Operations Research Society of America, Arlington, VA, 1983.

8. Bracken, J., *Lanchester Models of the Ardennes Campaign*, Naval Research Logistics, vol. 42 , 559-577, 1995.

9. Clemens, S. C., *The Application of Lanchester Models to the Battle of Kursk*, unpublished manuscript, 5 May 1997.

10. Hartley III, D. S., and Helmbold, R. L., *Validating Lanchester's Square and Other Attrition Models*, January 19,1993.

11. Hartley III, D. S., *A Mathematical Model of Attrition Data*, Center for Modeling, Simulation, and Gaming, Martin Marietta Energy Systems, Inc., Oak Ridge, Tennessee, 1995.

12. Kursk Operation Simulation And Validation Exercise – PhaseII (KOSAVE II), The US Army's Center for Strategy and Force Evaluation Study Report CAA-SR-98-7, September 1998.

13. Encyclopedia Britannica "City of Kursk." [http://www.britannica.com/bcom/eb/article/4/0,5716,47574+1,00.html]. January 2000.

14. Mapblast.com "Kursk." [http://www.mapblast.com/mblast/map.mb?loc=wo&CT=51.7328%3A36.1827%3A10000000&GAD1=&GAD2=&GAD3=Kursk&GAD4=Russia&W=425&H=250&IC=51.7328%3A36.1827%3A8%3A&lv=9&MA=1&zoom.x=46&zoom.y=29]. March 2000.

15. Mapblast.com "Kursk." [http://www.mapblast.com/mblast/map.mb?loc=wo&CT=51.7328%3A36.1827%3A400000&GAD1=&GAD2=&GAD3=Kursk&GAD4=Russia&W=425&H=250&IC=51.7328%3A36.1827%3A8%3A&COUNTRY=RUS&lv=7&MA=1&zoom.x=44&zoom.y=49]. March 2000.

16. Tsouras, P. G., *The Great Patriotic War*, Presidio Press, California, 1992.

17. Ziemke, E. F., *Army Historical Series Stalingrad to Berlin: The German Defeat in the East*, Office of the Chief of Military History, United States Army, Washington D.C., 1968.

18. Venables, W. N., and Ripley, B. D., *Statistics and Computing, Modern Applied Statistics with S-Plus*, Second Edition, New York, 1997.

19. Hamilton, L. C., *Regression With Graphics*, Duxbury Press, Belmont, California, 1992.

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center..2
   8725 John J. Kingman Rd., STE 0944
   Ft. Belvoir, Virginia 22060-6218

2. Genelkurmay Baskanligi..1
   Personel Baskanligi
   Bakanliklar
   Ankara, TURKEY

3. Kara Kuvvetleri Komutanligi..1
   Personel Daire Baskanligi
   Bakanliklar
   Ankara, TURKEY

4. Kara Kuvvetleri Komutanligi..1
   Kara Kuvvetleri Kutuphanesi
   Bakanliklar
   Ankara, TURKEY

5. Kara Harp Okulu Komutanligi...1
   Cumhuriyet Sitesi
   Kara Harp Okulu Kutuphanesi
   Bakanliklar
   Ankara, TURKEY

6. Dudley Knox Library...2
   Naval Postgraduate School
   411 Dyer Rd.
   Monterey, California 93943-5101

7. Professor Thomas W. Lucas, Code OR/Lt..2
   Department of Operations Research
   Naval Postgraduate School
   Monterey, CA 93943-5000

8. Mr. Ron Fricker...1
   RAND
   1700 Main Street, P.O.Box 2138
   Santa Monica, CA 90407-2138

9. Turker Turkes..1
   76 Sok. No:66 Daire:5
   Guzelyali 35290
   Izmir, TURKEY

10. Chairman, Code OR...1
Naval Postgraduate School
Monterey, CA 93940-5101

11. Mr. Walt Hollis...1
HQDA
ATTN: SAUS-OR
The Pentagon
Washington, DC 20310

12. Mr. E. B. Vandiver III...1
US Center for Army Analysis
6001 Goethals Road
Fort Belvoir, VA22060